

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACCTC
AGTGACTGAGGCCAGGCCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAACTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTTCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCTCGCCGCTGTG
CCTGCTATGTGCCACAGAGGTGCACTGTACATTTTCGGTACCTGACCTCCATCCCAG
ATGGCATCCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAACCTGGAGTTACTCATGCTGCACAGT
AATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTGCAGGTC
TAAAAAATGAGCTATAACAAAGTCCAAATCATTCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCCAAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCCTTTGCTATGGTACCATCTGGAGCTTTCCTATGTACAA
AGCCAACCATTTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAAGTCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAACCTGGAACGCACTGTCCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAACTGCAGATGGGCTGACAGCTTTGA
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
CAAGGTCATTACCAATGTGTGGCTGCCAACCCTCAGGGGGCCGACTTTTCCAGTTTT
AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
GTGGATCTGGACTTGGAGAACCCAACCTCCAGTGTTTCCCTTAAGCAGCCAGCATCTT
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
TACATAGGAAGAACAACATAGAGACTTAATACATCGGCGGCGTGTTGGGATTCCACG
CTCCGGCGATTACAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
GCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAAGTCA
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCTCCAGATGAAGAATTCATGGTTC
TGAAAATAAGGCTTCTGGTGTCCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
CACAAACACTACAATCTGAGCACCTTCTGATTTCAAATTATTTAGTGTAAACAAACG
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
ACCAACCAAAACCCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
CAGGCAGGAAGAGCATCTTCCCAAAGTGACACACCCTGTAACAGGGGGGAAACATGGC
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
GCAGACCTAGCAGTAGTGACATCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
GTCACCCTTCAGGTTACACACCACTGCCTCGTCTTTATTTACATTCTAGAAACAA
CAATACAGGTAACCTTCCCCTTGTCCAGGCACTTGGGAAGAGAGAGGACAATTTGGA
GCAGAGGGAGAGTTAAAAACCCACATAGAACCCCAAGTTCTCCGACGGCATAGACAC
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCC
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCCTTCCGCAGAGGGGGCTCACAG
TGGCTACTGCAGCACTGTCAGTTCCAAGTTCATCCACAGTGCCCTCCCCAAAATA
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
TCCGGAGGGGAAAGTAACCACGTGATTCTTACGGAAGCAAGCATGACTTCTGCTCC
AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
TGGGACCATCCAACTGGGAAAGATTGAGTGGAACAACCACTTCCCAGCCCCC
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACCTCCCTTG
CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAATAGCTCCTCTTTTACC
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCCGCCACCAGC
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
CTAGCACCATAAGCAAGAGATCTAATACATTAACTTCTTGTCAACGGAAACCCCCA
CAGTGACAAAGTCTACTGCTACTGCATCTGTCATTATGTCTGAAACCCAACGAACAA
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
TGATACCCCCCTTGGCTTTTCTAGTCATTCCCCACGACAAGATGATGGTGGAATGTAAG
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCTCCTCAGC
AGTGGGGGCGCCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
TTCTGTGCCAGCACTAAGGGTAGATAAAACCACAGAATTCTAAATGGAAGCCCTCTCC
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
AAAGGCCAGCAGTAAGCATGTCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGATAAGAAACC
TGGTCAAACCCAACCTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
AGACGTTTTTCTTCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGTCTGTG
GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
GGAAGTACTGTGGAACATAAGTGCAGAGTGGAGGGTATGCCGAGGGCCTACGGTTTC
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACTGTTGGTTAAGA
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGGCAAGCCATCGTTG
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
GTGTTCACTGGGTCCCTTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT
CCAGATTTTTCTTGTATCCAAATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGT
GAGGGGCACCTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAACTGCCTCTCAG
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCAGT
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

Figure 2

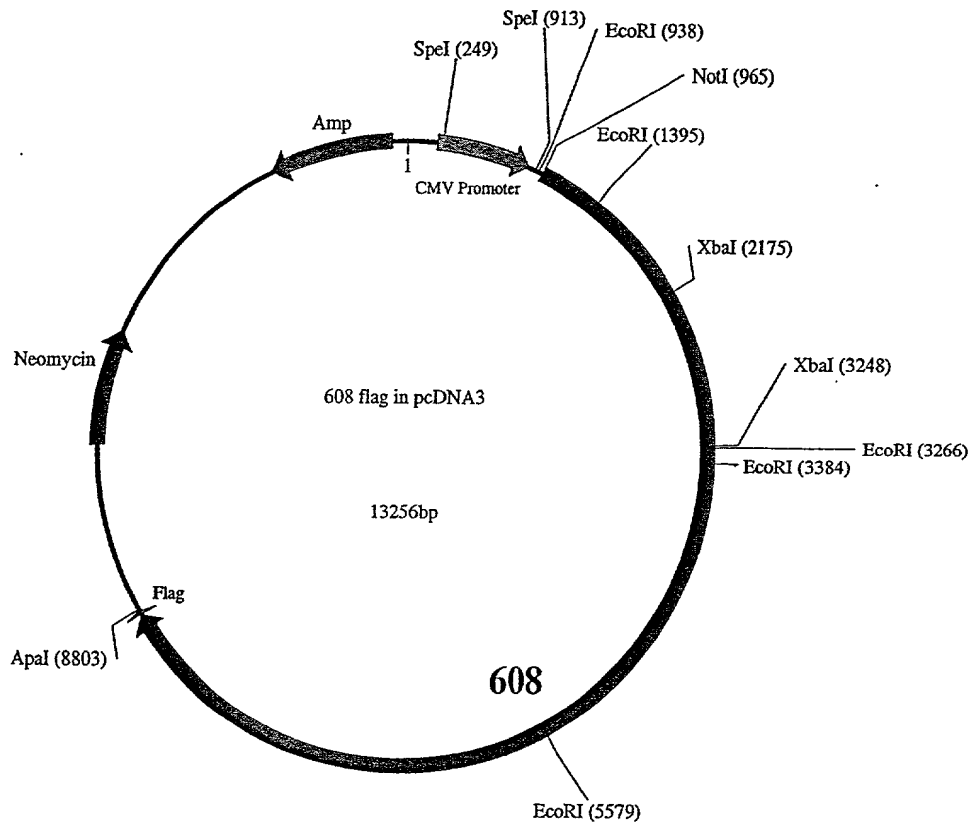


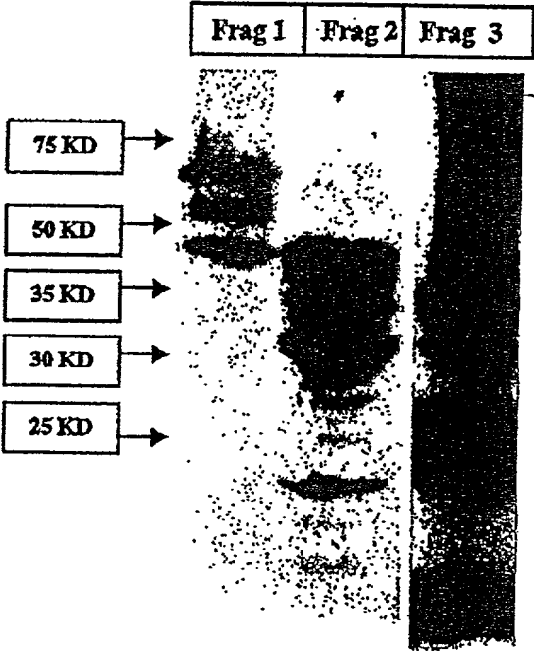
Figure 3

MQVRGREVSGLLISLTA VCLVVT PGSRACPRRCACYVPTEVHCTFRYLT SIPDGIPANVE
RINLGYN SLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNGS
ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
ADPFWFQQEKIVLQLNRTATTLSLTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNNPK
LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
DSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKL SASALTGSEAGKQVSGVHRK
NKHRLIHRRRGDSTLRRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
KPVPLAVPLVELTDEEKDASGMIPPDEEFMVLKTKASGVPGRSPTADSGPVNHGFMTSI
ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIFPSVAE
IRDSAQAGRASSQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTESYGPQIPITGV
SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
KNPHRTPVLRHRHRTVRPAIKGPANKNVSVQVPATEYPGMCHTCPSAEGTLVATAALS
VPSSSHSALPKTNVGVIAEESTTVVKKPLLLFKDKQNV DIEITTTTKYSGGESNHVIPTE
ASMTSAPTSVSLGKSPVDNSGHL SMPGTIQTGKDSVETPLPSPLSTPSIPTSTKFSKRKTP
LHQIFVNNQKKEGMLKNPYQFGLQKNPA AKLPKIAPLLPTGQSSPSDSTLLTSPPPALST
TMAATQNKGT EVVSGARSLSAGKKQPFTNSSPVLPTISKRSNTLNFLSTETPTVTSPTAT
ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALTADTPLAFSHSP
RQDDGGNVSAVAYHSTTSLLAITELFEKYTQTLGNTTALETTLLSKSQESTTVKRASDTP
PPLLSSGAPPVPTSPPPFTKG VVTDSKVTS AFQMTSNRVVTIYESSRHNTDLQQPSAEAS
PNPEIITGTTDSPSNLFPSTSVPALRVDPQNSKWKPSWPPEHKYQLKSYSETIEKGK RPA
VSMSPHLSLPEASTHASHWNTQKHA EKSVFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
GKAASFTVPANS DVFLPCEAVGDPLPIHWTRVSSGXEISQGTQKSRFHVLPNGTLSIQRV
SIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
RPTVSWILANQTVVSETAKGSRKVWVTPDGTLLIYNLSLYDRGFYKCVASNPSGQD SLL
VKIQVITAPPVIEQKRQAIVGV LGGSLKL PCTAKGTPQPSVHWVLYDGT ELKPLQLTHS
RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVVILTVEEGETIPRIETASQKWTEVN
LGEKLLLNCSATGDPKPRIIWRLPSKAVIDQWHRMGSRHVYPNGSLVVGSVTEKDAGD
YLCVARNKMGDDLVL MHVRLRLTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVPEV
SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFN NVGMAEEGDYICSAQNTLGK
DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS
NDRFTFHANRTL SIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

DSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLLYKATR NKSGKYRCAARNKVG YIEKLIL
LEIGQKPVILTYEPGMVKS VSGEPLSLHCVSDGIPKPNVKWTT PGGHVIDRPQVDGKYIL
HENGTLVIKATTAHDQGN YICRAQNSVGQAVISVSVMVVA YPPRIINYLP RNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLLGTDYATTYIQVL (SEQ ID NO: 2)

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4



9/90
Figure 5

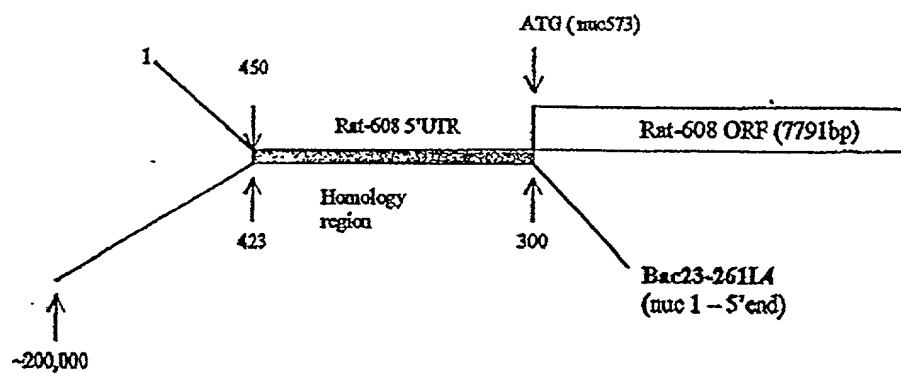


Figure 6

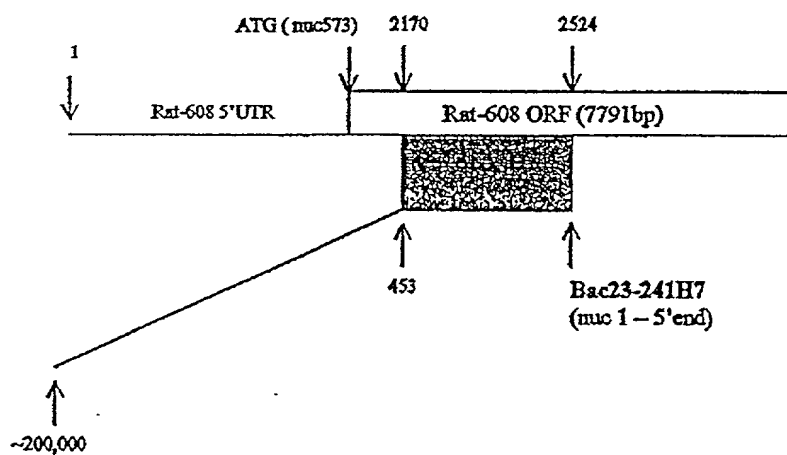


Figure 7

Nuc 1

↓
TTTGGAAACCAACCCAGATGCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCATTTA
TCCAATGGAATACTACTCAACTTATTAACCAACGACTTTCATAAAATTTTAGGCAAAT
GNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACACTCATGGTATGC
ACTCACTGATAAGTGGCTATTGTCTATGGAGTGATTTAAAGGGAAGAACACATAGC
TTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGTTTGGCTTATGAGTGAATCT
TGTTTCAGTGGATTACCGGAAGAAATAATAAGTGAAGTGTAGGAAGAAGTAGTTAATCA

↓ Nuc 390

AGGTGACAAAGTATCCTGACACATTGGGAAAAGACCACAGTCCAGGAAACTGAGTCTTAA
GGATTTCATATTAACCTCCAGTTCCCATGTGCAGCTCTGAGACTTTGGCAGATCAGACACTT
AACTTACAGCTTCTACACAGAGCAGTTACTATCCTTGCATTCACAGATGGAGTGTGA
CCATTAAAGCTGACCTGAAACATGAGTCTGACTTGITAATAATCTTAAATACAAATGTGT
TGTAAGTATGTGACCAAGAGCATGGTCACTGCTATTAACTTTGATGTTCTATGGACTCT
TAATTTTATGGTAGAAATGTCAACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAAGAG
GATGGCTTTCAAGTTTGAAGTAAATTCAGTATGTGTTCTGGCATCCCTTTCTAAAGCA
ATTTAAACCCCAAGTAGGCATAATTTAATGCTTACTTCATCAGAAATATGCTAATTGAC
TCTTCTAAAAAGACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGCATAAT
AAATAGGAGAAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTT
ATAGGTTATATAAGAGGTGGGAAAGGAATATTGAAACATCTAGAAGTAAATGTATCCTG
AGTAGCGATCCTGGGAAAATACGTAAGTCTAACACACTGCAATCATCTCTGTGTTTGGT
GGAGCTGAGGTCTGGAAGGCTGACCTTGGTTAGAAATAACCTACCGAATACAGAGCTAT
GACGTTAGTCTGGAAGGAGCTTTGGAAAGATGACAAGCTGTAGCTGGCCAGAACATCTA
GATGCCATATTTCAGGCAAGTGTCCACATGCGGACATCTTAAAGAAATATGGTTGTCTG
CAGTGTCTAAGGACCTTGTCTGTCACACAGGTCTCCAGGTTAGTGTCTAAGTCTGATGCT
TIGACTCTTAAATTTCTAAGTATGATCATTAAATGACTAGAAATCACTTGGTGATTAGCAACTG
GATATGGAATATTACTAATTTGTACCAAGCCAGGCCACCTCAGCTTTGGCAGCTCCATTG
ATTGCTGGAGGCCAGTCAAGTGGGTTTGAATCAACTGTACTGTTTCTACTTACAAGAGCC
ATTACCTGAGATGAGTCAATTTTCTTCAACAAGTCTTTTGAAGAGTCAATTAGACATATT
CTGATGAAGTAAAGCATATAAAGTGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGA
TGCTATTATAATGTGGAATAAACTGACTTTAAAAAAGGCTTATGATACTTGTGACA
GAGTAAATCTTCCATAAATATCATCTGCAATTTATAAATTTTTCATAATCCATCAATTA
AAACCTTTAGAAATTTTGTAAACACAAAGATCCCTAGGCCCTGCCCCTAGGATGGTCTGTA
TGGTGGGCTGAGAGATGGAGCTTAAAGAACTTACTGTCTCAGGAGCAGATCTTCAGAAC
ATCTGGCTCAAAACATTTATCCAAATGCTCATCAAGGCTCACTCAGATGTGCTTCAACC
ACAGGGATTAAACAGTCATTTTGTGACATTTCTCAAAAGGCTGGAAGGCTGTAGAGGAA
CAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTGATAACCTTCACCAACA
AACAACAACGACAACAAACCCATAAAATTTATCAAGGCAAAATGAACTAAGCCATATGCA
GAAAAAGTATTATATGTTCTCATTTGTGGGGTGTTTTCTTAAATAGTCAAAATATGAGAAT
ATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTTGTAGATTAGAGGACA
CAAAGCAACAACCTACAGAGTGAAGTAATCCAGAGACTTAATGTATAATATGAGGACTGTA
TTAATAATTTCTATTAAAGATACACAGCAAACGAGTGTATCTTACTAACACACACACTTAC
ATAGAGAGAATAAAGTGATAGATACGTTTGTATTATCTTCATGTAGCTGATAATTTCATAT
TGTACACCTCAACATAGATAACCAACAAGAGGAAGAGGATAGGTGGCTCTCCAGGGC
GGAAGAGTACATTGGAAGTCAAGACACCAATTGTGTAGATGTACCACATGGAGGAGCTAGA
GAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCCCTATAGGTGGAACAACATTTAGGCT
AACCAGTACCCCGAGCTCTTGACTTAGCTGCATATATCAAAAAGATGGCCTAATCGG
CCATCACTGGAAGAGAGGCCCATTTGGACTTGCAAACTTTATATGCCCCAGTACAGGGGA
ATACCAGGGCCAAAAAGGGGAGTGGGTGGGACGGGAGTGGGGGTGGGTGGATATGGG
GGACTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATACTAATAAAAAATGAAAAA

↑ Nuc 2817

AAAAAAAAAAAAAAAAAAAAAGGAAGGTGAGACACCTCACTTCACTGCTATCTCAACTT
GCAACAGAAAGGGGAGTCACAAACCCAGGACAAACACAGTGATTGAAGCGTCTTTGAA
TGTTATTGCTGTTGTTGTTACCAATCATTAGCATATATTCATTGTGAAAACCTAACGGGGT
CTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGCACCACCTACCA

09991630-030402

Country	Year	Population (millions)	Urban population (millions)	Urban population (%)	Population density (per sq km)	Urban population density (per sq km)	Population growth rate (%)	Urban population growth rate (%)	Population growth rate (%)	Urban population growth rate (%)	Population growth rate (%)	Urban population growth rate (%)
Algeria	1980	10.0	4.0	40.0	100	250	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1985	10.5	4.5	42.9	105	260	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1990	11.0	5.0	45.5	110	270	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1995	11.5	5.5	47.8	115	280	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2000	12.0	6.0	50.0	120	290	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2005	12.5	6.5	52.0	125	300	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2010	13.0	7.0	53.8	130	310	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2015	13.5	7.5	55.6	135	320	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2020	14.0	8.0	57.1	140	330	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2025	14.5	8.5	58.6	145	340	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2030	15.0	9.0	60.0	150	350	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2035	15.5	9.5	61.3	155	360	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2040	16.0	10.0	62.5	160	370	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2045	16.5	10.5	63.6	165	380	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2050	17.0	11.0	64.7	170	390	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2055	17.5	11.5	65.7	175	400	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2060	18.0	12.0	66.7	180	410	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2065	18.5	12.5	67.6	185	420	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2070	19.0	13.0	68.4	190	430	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2075	19.5	13.5	69.2	195	440	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2080	20.0	14.0	70.0	200	450	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2085	20.5	14.5	70.7	205	460	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2090	21.0	15.0	71.4	210	470	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2095	21.5	15.5	72.1	215	480	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2100	22.0	16.0	72.7	220	490	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2105	22.5	16.5	73.3	225	500	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2110	23.0	17.0	73.9	230	510	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2115	23.5	17.5	74.5	235	520	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2120	24.0	18.0	75.0	240	530	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2125	24.5	18.5	75.5	245	540	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2130	25.0	19.0	76.0	250	550	1.5	2.5	1.5	2.5	1.5	

▼ Exon 1 (Nuc 3661)

↑ Nuc 4203

TGTTAAATACAGTTTGAGGGCTTAAGTGTACGGGAACCTCATGTGGTATTCATTAAACGGCTCT
CTTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTCTTCTGTTTCCAGCTACCTTGTAC
CATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATCAAGCAGAGAGTATT
CAAAAAATATTCAGTGATGTAAACAGTGACAGCTGTAGGCATAGAAGTAATCATTAGTAAATC
TTAATTTGGGTTTAAACTCATTCTATAACAGCTCCAGGTTGGGAGGGGATCTAGTGAGCCCTTCGC
CACGTGCGGGTTAAAGATATTTTCTAAACAAGAGAAGCAGAATTCTTCTCTTGGCCATGCTCC
CCATCACTGTGTCAAGTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAGAGCAGACAGTGTTA
TGCCGTGCAAAGTCAGAGACTCAGCCCTCCAGCTGCTCAGTTTACTGTCTCTCCGGTCATT
AGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGCAGACATGCTAGAAA
GAAATTTGACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAAGTATTTTGGAAAGGTTAAG
CCAAATAAATAAAACACATATTAAATAATACAATGTTACAAAATTTGATCATATAAAGAA
GTACATTTCATAAATGCAATTGTGAAAAATATATATAAATTTTTATCTATTTTACTGGTGCAAAG
TTTTCTAAATTGCGACATGTACTATTTTTATATTATAAAAAATTTTAAAAATGTATATAAA
AGTGTA AAAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAAACHTTAATTGTCCCGAT
AACATTATTATGATCTCTAATGACAGGGATCTGCTTTTCATTGGGAAATGAGAAGCTATG
AAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCAATGGGAAGCTAGCACAC
ACTGGTTTATAAAGAGAACAGTTTCTCTGAGTCTATGCAAGTTTACACTCTAGGGAATAAG
AGTTCTCTTTCTCCAGATTTCCTACTAGCATTGTTGTCATCATTTATCTCTTGATGATGAG
CATTATAAGTGGAATAAGATAGGATCTAAAGGAATGTCAATTTGGATGCGCTGAACAAAT
CTTTACGGTCTTTCTTTTCAGTTCACTAGTCTATTTCATTTATTGGATAATTGGGGGATGGTGT
TAATTTTTTTGCAAGTTCTTAATGGAATTCAAAAAAACAAAAACAAACAAACAAAAA
ACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACA
GAATTTATTGCTACATGTTCTATTATTCACGACGCCAAGAATGTGGAACCAACTTACGTA
GCGGTCAAAATATGACCGGATAAGAAAAATGTGGAATGTGTACAACAGAGATCCCATGTG
GCCATAAAAGAGTGAAATCATGACATATGACAGGAATGGATGCAAGCTGGAATCAATTG
GGCTAATCAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAAGCTTCTCTGACAAAC
GAAGCTAGATTACACTTGTACGTGGCATGTGTGTTTAGAATTTTATTTFAGTTATACACT
ATTCTAATCTGTGAGTGTTGTATAAAGGCATGCATGTAAAGCAAAAAACAAGCTAGCTGGGG
TGGGTAGGAGAGAAAGCAATGAGAGGAGTTAAATAAGAACGAAGCATAGTAACATAGGTG
CCAGGATGAAATGCAATTAATTTGATGCTAACTAAACCAACAGACAGGAGGCCACAGTTTCA
AACCAGGGTGAAATCCACGACAGAGAAAGGGGAAGTAGACACAAAGTTTCCCACTAAC
CAAGAAGCCATTGCAAGTTGCTGCCTGCTGGGAGGGCGCTCCAGTTTCTCCAGTCTGAC
ACTGTGTATAACAACCAAGTTGACAATACAAAGTTGGCATGATGGAATGGTTTTTGTGCTATT
TTTCAATTTTTTCTTACTGTTTTGTTGTTGTGTTGTTGTTGTTGGTGGTGGCTGTGTTTTC
ATTGTTTTCTTTTGAGAGAGAGAGAAAGAAACATGAAATTTGGTGGGTAGGAAGCTGGAACCG
ATCTGGAAGAAAGTTGGGGAAAGAGAAAAATTTGATGAGCATAATTTAAACAAACAAACA
AACAAACAAAGGTTTCATTTTGCCACAAAAAGGTGTGAATTAATTAACAGTTTACGACT
CTTAAAGAAAAATATCCCAATTATTCACAGAGTTGCTATGTATGCTGTGCGCTAGGACTTTG

↓ Exon 2 (Nuc 6559)

CTTGAACITGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGAA
GTCAGCTGCTTGCCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCGGGAGCAGGGTC
TGTCCTCGCGATGTGCCTGCTATGTGCCACAGAGGTGCACGTGTACATTTGGGGACCTGAOC
TCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGT

↑ Nuc 6755

GGACCTTGCTGATCTCCTTCTCAGAGAGGGACCACCTGATTTTCTGGTACTTTGCCCCC
AAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTTCATACAAACCTTATATTG
TGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAAGCTTGAAGCCCTCTTAGIT
GTTTGGGAGCCTCATAATTATGGGGCAGCTTTATCTGAATGAATTTAAATGAAAAAGAT
ACAGTTTTCTGTTAAACAATCATTATGATACCAAGGAAGAGGAATTGTTCATTGAATATTTAA
AAAAGCATTCTTTTGAATTTATAAATACCCATTACAAAATGGCTTACTTAAAATACTTG
CCTTACTAAATCTGACAAATTATGGTGATATTTTGAAGGTTTATGAAAATTTGTTTATGTGT
ATAAATGCACAAGAAATGGGATATGCCATCACCTATGTGCCATTAGTGAGCATGTACAGT
ATGCCAAAACACTATTGTTACAGTTTGGAGGAAGTAATGGGGGTGGGGGAGCAACAAGGGT
TATAACCGTATACCCAGTGCTTGGGAAGCGATTGCAAACAGTAAAGACTGACATTGTGT
CTCCCTATGAGGGAGGGGCCCTTGGGCTGAGCACTTTGCAATGAGCATTGTCTCATTTGTGCT
GGCAGGTTTTATGATAACTTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAAGCTTCAACT
GAGAACATGCCCTGAAGAAGATCAGATTATAGGCAGGCTGTGGGGCATTCTTAAATTAG
TGATTCATGGGGCAGGGGCCAGTCCATTGTTCTGTGTACCATTTCTCAGGCACATTAAAAA
AAAAAAAACAGGCTGAGCAAGTGTCAAGGAGCAAGTCACTGAGCAGCAGCCCTAATGA
TCTCTGCATCAGCTCCTCCAGTTCTACCTATTGAGTTCCTGTCTAGCTCCCTA
CAGTGATGAACAATGTGTGGAAGTATAAGCCAAAATAAATCCTTTCTTCCCCAACTGTGCT
TTGGTCATGATGTTTCATCACAGTGATAATAGTCTCATGAAGATGCTGTGTTTATAACA
CCTTTGGACTAAATCTGTGTATCTATAGCTGAGGAAAATGGAGCATAGAAAGTCTCCAGA
CTACACCAGAGTGTAATCTGGGCTGAGCTTAGAATCACACCCACGTGCACCTCCACTGCC
GGGGCTTCTTAAOCGGAACACAGTTGTAAAAGGGAATTTTCTGTTTGTTCATTTTGACA

Exon 3 (Nuc 8089) ↓

TGTGGACTTTAATTGACGATTTCATCTGAAGCTGAAAATGATTTTTTCCAGGTATAACAGC
CTCACTAGATTGACAGAAAATGACTTTTCTGGCCTGAGCAGACTGGAGTACTCATGCTGCACA
GCAATGGCATTTCACAGAGTCAGTGACAAGACCTTCTGGGGCTGTCAGTCTTGCAG

Nuc 8218 ↑

GTGAGATAGGTAGAGGGTGATGGAGGCTGAGAAGAGAGGTTGCAACTGTGGGTTATACCC
AAAAGCTGCTGATTCCCGTGGGAGACATTCTATAAGCATTCTATAAACTAGAGGCAGATA
TCAAGGAAGGATTTCAATTGTAATGCAATTTATGAGAAAATTTGAATATTAAGAAAATG
CTGGGGAATAATGCTTACACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATA
AAAGCAGCATGCGGCGAGACCTGCAATTCCTGTCCCTGGAAGCACTGTTGAGAAAT
CCCAGAGACTCATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAA
ACTTGACTCAGAACTAATGTGGAAGCATCAGGAAGACAGCCAACATCTGCTCTACT
CATGCATGAATAAGGGATCCAGAGAGAAGGGAAGAAAAGGAAGGAAGGAAGGAAGG
AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
GGAGAAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
TTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGTGGGATCCCTTATAT
AAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAAATTTTGAATTTTCTCTAC
TTTTATCCTCTACCATCTTCAAACTGAAACTACAATTTATCCACAAATGAAGAAAATGC
TGTAAAGAGTTTTCACACACCGAAGTGGGAACTTAAGGATTAGACAAAGTCTAACAATGAG
AATGGGGAGAACAAAAAGAGACTGCACAGGGAGCCCTTCTCTGCTTATAATCTTGACAC
TTGAGAAGCTAATTGACGCTGCATGACTACTCACTCTTTAAGCAAACAATGCTGTTGTTT
ATGAAAAGCACAATAAAGTACATATGTCCATAATATTCATCAAAATTTGCATGCAGCAC
ATAATAGCAATCAAAGCAATAACACCCACTGTTCCACAGAGACTTTAAACATGAAACTGGA
ACTATGTCTAGTGTTTTGAATTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTG
ATTAGGTCATCAGACAGCATTTTGAACATGTATCTTCAGGAGGAATCAATCATGTATCT
GCATGAAATTTCTCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTGATGGAGAAACATT
GGGTTTGAGGTTTTACTCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGG
CTTTTATTTATGAGGGATGTTGGTATTCCAGAAAATTTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAAACAGAAAAATATGTAAAAAGTTATTGTTTTATTAGTATTTTCATGTTCTTTCTT
TTTTAAAAATGGTATGCTTAGAAGCTAATTAAGATTAGATTAGATTAGATTAGAAAATAATC
AGAGAGGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTTTTGCTTCTAATT
CAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAGATGTGTGCCAGCTGAG
GAGTGATAAACTGTGTAACTGAGTGCTATGTAGAAAACAGAAAAGGAAGTGAAGGGTTGA

TGTGCGCTGCAACATCTTGAAAACATTGCGGTACATGATGGAAGCCAGGCACAAAAAGCC
ACATATTGCATGGTTATGTTTATATGAAATGTTTAAAAATACATGGATTCTTAGCAACAGA
GTAAGATGTTACTTAGGGTCAGGAAAAGATTAAAAAATACTATTGATGTGGAATG
ATCTTAATTTGGGGAAAAGACAATTTCTAAGACGAAATAGTTGAGGTAGATATAGTTAT
ATCCCTGTGGATATTGTAATAAACAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGG
GCAGGAGGAAAGTGAAATGAGATGGTAGAAAGGAAAGTCATATAACCATGGCTTCTCTCGTG
GGTGGAATCTAGATATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGGAAGGATC
AAAAACCAACAGGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAAGTTCATGGTCAA
TGTGTGTGGAGACACCATAATAAAACTCCTTTTTTGTGCTAACTAAAAACCACTAAAAATC
TAAAAACAAAACATTTTTGACACAAGAATTATTATTATTCAATAAAGATGTTTAAATGGGG
GAAGTTGAAGTTTCATTGATAGTCTCATAAATCTTAAATGTATTTAAACTGCTTTTTACGTTT
TTTATTATTAACTACTCTGTCTGTCATTATTATCATCATCATTATCGTCATCATCACTA
ATGCTTTTCAACCATACACAAATGTAGGCAGAAAGAGTGTAATCCACTTAGTGAGGGCACTT
GGAAGAGGGAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAGGACAGTGAGAGGGAA
ATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCATAGTCCACTACATTACTTTG
TATTCTAATATTAAGAAAATAATAAACCCATTTCTGTGCACTTATCACCCAGGCTCAACAG
TTATCTTGGCCACAGATCCTGTCTCACTGCATCCTGTGCCACCTGAGTCCACTTAGCGTTCTG
AATCCAATCCAGGGCATGATGCTTACTCTACACAGAACTAAAGATTAAAGAGAGTTTAA
AAGTAACCATGACATCTCTCTGTTCTTTAGCGATAAGTTCTTAATATTATGCGCTGCTTGT
GTATGTTCTAATTTCTCTAATATTGTGCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGT
TGGAGTTAAGGTCCCATAGGATTAATCTCAACATATTCTATATTTATAAACTTTCTCTCT

Exon 4 ? (Nuc 11286) ↓

TTGTGAAAGTTCCTTTGAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATG
AGCTATAACAAAGTCCAAATAATTGAGAAGGATACCTTTGTATGGAATCAGGAGCTTGACCCGG
TTGCACCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTACGGACTCACCTTGC
TCCGCTTGCTACATCTAGAAAGAAACCGGCTGACAAAAGCTCCATCCAGACACATTTGTCTCTT
GAGCTATCTCCAGATATTTAAAACTCCTTCATTAAAGNAOCTGTACTTGTATGATAACTTCATTG
ACCTCCCTCCAAAAGAAATGGTCTCTCTCTATGCCAAACCTAGAAAGCCTTTACTTGCAATGGAA
ACCCATGGACCTGTGACTGCCATTTAAAGTGGTTGTCCGAGTGGATGCAGGGGAAACCCAGGTA

Nuc 11680 ↑

ACTATCTGTGTTGTTTGTCTTTTTTATARKACGTTATTTCTCAATTTTCAATTTAGAAATGA
TATCCAAAAGTCCCCATAAOCCTCCCCCCTTCCCTACCTACCCATTCOCATTTTTTGG
CCCTGGCATTCCCTGTACTGGGGCATATAAAGTTTGGGTGTCCAATGGACCTCTCTTCC
AGTGATGGCCAACTAGGCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGG
TACTGGTTAGTTCATAATGTTGTTGCACCTACAGGGTTGAA

Nuc 11967 ↑

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

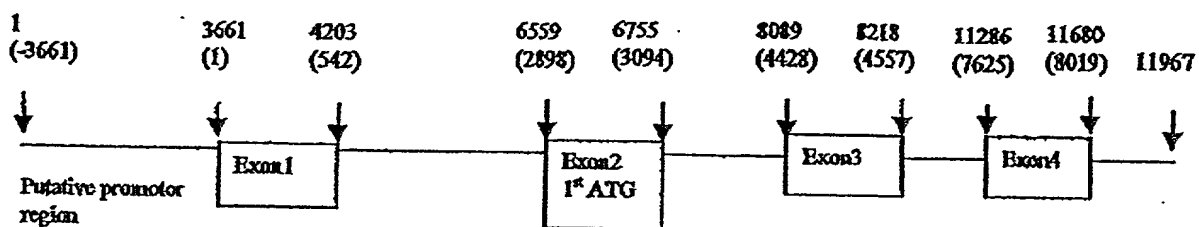


Figure 10

cDNA_rat	1	cgagagacgacagaagggttaoggctgcgagaagacgacagaaggggtccag
genomic_hu	1	-----
cDNA_rat	51	aaaaaggaaagtgtctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu	1	-----
cDNA_rat	101	atgtcacttcagtgtactgaggccaggcaaaacgogcggaaggattttgt
genomic_hu	1	-----
cDNA_rat	151	gtagcttgggacccctttcatagacactgatgacacgtttaogcaaaatag
genomic_hu	1	-----
cDNA_rat	201	aaatttgaggagaaaacgcctgggccttcggaaggagtgattgattagta
genomic_hu	1	-----
cDNA_rat	251	cttgcaagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu	1	-----
cDNA_rat	301	gggaggaggaagagcattacagagtttccagcagcagcaggaaagctttg
genomic_hu	1	-----
cDNA_rat	351	gttaatttggaaatggatgatagcattaaaataacagaagcgccctccagg
genomic_hu	1	-----
cDNA_rat	401	tototgaagottoagtcccccagctgaaagccagaaaagaactaagccaac
genomic_hu	1	-----
cDNA_rat	451	taagccttttgatccctttggaagcaaagaactttccttccttggggtga
genomic_hu	1	-----
cDNA_rat	501	agactotcctcagaagatttcctgtctctgcctatgttacaagaggaatc
genomic_hu	1	-----
cDNA_rat	551	aaaaccaagacagaagagctcaggatgcagggtgagaggcagggaagtca
genomic_hu	1	-----
cDNA_rat	601	oggcttgttgatotccctoaotgotgtctgcctgggtggtoacccctggga
genomic_hu	1	-----
cDNA_rat	651	gcagggcctgtcctgcgcgtgtgcctgtatgtgccacagaggtgcac
genomic_hu	1	-----
cDNA_rat	701	tgtacatttcggtaacctgacctccatccagatggcatccggccaatgt
genomic_hu	1	-----
cDNA_rat	751	ggaacgaataaaatttaggatataacagccttactagattgacagaaaacg
genomic_hu	1	-----
cDNA_rat	801	actttgatggcctgagcaaaactggagttactcatgotgcacagtaatggc
genomic_hu	1	-----
cDNA_rat	851	attcacagagtcagtgacaagaccttctcgggcttgcagtccttgaggt
genomic_hu	1	-----
cDNA_rat	901	cttaaaaatgagctataacaaagtccaaatoattoggaaggatactttct
genomic_hu	1	-----

cDNA_rat genomic_hu	951 1	acggactcgggagcttggtccggttgacactggatcacacaacacattgaa -----
cDNA_rat genomic_hu	1001 1	ttcatcaaccctgaggccttttatggacttacctcgctccgcttggtaca -----
cDNA_rat genomic_hu	1051 1	tttagaaggaaaacgggtcacaaagctccatccagacacatttgtotoat -----
cDNA_rat genomic_hu	1101 1	taagotatctccagatatttaaaacctctttcattaagtacctgttcttg -----
cDNA_rat genomic_hu	1151 1	tctgataacttctcgacctccctcccaaaagaaatggtctcctacatgcc -----
cDNA_rat genomic_hu	1201 1	aaacotagaagaacctgtatttgcattggaacccatggacctgtgactgcc -----
cDNA_rat genomic_hu	1251 1	atttaaagtgggttgtotgagtggtatgcagggaacccagatataataaaa -----
cDNA_rat genomic_hu	1301 1	tgcaagaagaacagaaagotcttccagtcctcagcaatgtcccctttgcat -----
cDNA_rat genomic_hu	1351 1	gaaccccaggatctctaaaggcagaccctttgctatggtacoatctggag -----
cDNA_rat genomic_hu	1401 1	ctttcctatgtacaagaaccaaccattgatccatcactgaagtoaaagagc -----
cDNA_rat genomic_hu	1451 1	ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt -----
cDNA_rat genomic_hu	1501 1	catagaacocctttggotccttgtctttgaacatgacanannntctggaa -----
cDNA_rat genomic_hu	1551 1	ataaggccgacatggctctgtagtatocaaaagccatcaaggacatcacca -----
cDNA_rat genomic_hu	1601 1	actgcattcactgaagaaaatgactacatcatgtataaatgcgtcattttc -----
cDNA_rat genomic_hu	1651 1	cacaaatotttgtgtgcagtgtagattataatcacatccagccagtgtygc -----
cDNA_rat genomic_hu	1701 1	aaottctggtctttatacagtgactctcctctgatactagaaggaagccc -----
cDNA_rat genomic_hu	1751 1	cagcttaccgagactccttcactgtcttctagatataaacaggtggtot -----
cDNA_rat genomic_hu	1801 1	taggctgaagacatttttaccagcatagaggctgatgtcagagcagacc -----
cDNA_rat genomic_hu	1851 1	ctttttggttccaacaagaaaaaattgtcttgacgtgaacagaactgcc -----
cDNA_rat genomic_hu	1901 1	accacacttagoacattacagatocagttttccactgatgtcfaatccg -----
cDNA_rat genomic_hu	1951 1	tttaccaagggcggagatgagagcggagagactcaaattggaacctgatoc -----

cDNA_rat	2001	tgatgatgaacaatoccaaactggaacgcactgtctctgggtggcggaact
genomic_hu	1	-----
cDNA_rat	2051	attgccotgagctgtccaggcgaaggcgacccttcacctcacttggaaatg
genomic_hu	1	-----
cDNA_rat	2101	gottotagctgatgggagtaaagtgagagcccttacgttagcagaggatg
genomic_hu	1	-----
cDNA_rat	2151	ggcgaatcctaataagacaaaaatggsagttggaactgcagatggctgac
genomic_hu	1	-----
cDNA_rat	2201	agctttgatgcaggtctttaccactgcataagcaccatgatgcagatgc
genomic_hu	1	-----
cDNA_rat	2251	ggatgtttcacatacaggataactgtggtagagccctatggagaaagca
genomic_hu	1	-----
cDNA_rat	2301	cacatgacagtggagtcacgcacacagtggttacgggtgagacgctcgac
genomic_hu	1	-----
cDNA_rat	2351	cttccatgcctttccacgggtgtccagatgottotattagctggattct
genomic_hu	1	-----
cDNA_rat	2401	tccagggaacactgtgttctctcagocataagagacaggcaaatcttta
genomic_hu	1	-----
cDNA_rat	2451	acaatgggaccttaagaatattacaggttacgccaaaagatoaaggtcac
genomic_hu	1	-----
cDNA_rat	2501	taccaatgtgtggctgccaaacccatcaggggcgacttttccagttttaa
genomic_hu	1	-----
cDNA_rat	2551	agtttcagttcaaaagaaaggccaaaggatggttgagcatgacaggaggg
genomic_hu	1	-----
cDNA_rat	2601	caggtggatctggacttggagaacccaaactccagtgtttcccttaagcag
genomic_hu	1	-----
cDNA_rat	2651	ccagcatctttgaaactctctgcatcagctttgacagggtcagaggctgg
genomic_hu	1	-----
cDNA_rat	2701	aaaacaagtctccggtgtacataggaagaacaacatagagacttaatac
genomic_hu	1	-----
cDNA_rat	2751	atcgccggcggtggggattccacgctccggcgattcaggagcataggagg
genomic_hu	1	-----
cDNA_rat	2801	cagctccctctctctgctcggagaattgaccgcgaacgctgggcagcact
genomic_hu	1	-----
cDNA_rat	2851	tctagaaaaagccaaaagaattctgtgccaaaaagcaagaaatacca
genomic_hu	1	-----
cDNA_rat	2901	cagtaaagccagtgcactgggtgttccctcgtggaactcactgaogag
genomic_hu	1	-----
cDNA_rat	2951	gaaaaggatgcctctggoatgattctccagatgaagaattcatggttct
genomic_hu	1	-----

cDNA_rat	3001	gaaaactaaggcttctggtgtccaggaaggtcaccaactgctgactotg
genomic_hu	1	-----
cDNA_rat	3051	gaccagtaaatcatggttttatgaagagtatagcttctggcxcagaagtc
genomic_hu	1	-----
cDNA_rat	3101	tcaactgtgaatocacaaacactacaatctgagcaccttootgatttcaa
genomic_hu	1	-----
cDNA_rat	3151	attattttagtgaacaaacggtacagctgtgacaaagagtatgaaccocat
genomic_hu	1	-----
cDNA_rat	3201	ccatagcaagcaaaatagaagatacaacccaacccaaacccaatcattatc
genomic_hu	1	-----
cDNA_rat	3251	ttccatcagtagctgaaattcgagattctgtcaggcagggaagagagcat
genomic_hu	1	-----
cDNA_rat	3301	ttcccaagtgacacccctgtaacagggggaaacatggctacctatggcc
genomic_hu	1	-----
cDNA_rat	3351	ataccoacacatatagtagctttaccagcaaagccagtagcagctcttgag
genomic_hu	1	-----
cDNA_rat	3401	ccaataaatccaacagaaagttatggacctcagatacctattacaggagt
genomic_hu	1	-----
cDNA_rat	3451	cagcagacctagcagtagtgacatctcttctcacactactgcagacccta
genomic_hu	1	-----
cDNA_rat	3501	gcttctccagtoacccttcaggttcacacaccactgcctogtctttattt
genomic_hu	1	-----
cDNA_rat	3551	cacattccctagaaacaacaatacaggttaacttcccttgtccaggcaactt
genomic_hu	1	-----
cDNA_rat	3601	gggaagagagaggacaatttggagcagagggagaggttaaaaaccacata
genomic_hu	1	-----
cDNA_rat	3651	gaaccccagttctccgacggcatagacacaggactgtgaggccagcaatc
genomic_hu	1	-----
cDNA_rat	3701	aagggaactgtctaacaaaatgtgagccaagttccagccacagagtacc
genomic_hu	1	-----
cDNA_rat	3751	tgggatgtgccacacatgtccttccgagaggggctcacagtggotactg
genomic_hu	1	-----
cDNA_rat	3801	cagcaactgtcagttccaagttcatcccaagtgccctccccaaaactaat
genomic_hu	1	-----
cDNA_rat	3851	aatgttggggtcatagcagaagagtctaccactgtggtcaagaaacoact
genomic_hu	1	-----
cDNA_rat	3901	gttactatttaagggaacaaacaaatgtagatattgagataataacaacca
genomic_hu	1	-----
cDNA_rat	3951	ctacaaaatattccggaggggaaagtaaccaagtgatttoctacggaagca
genomic_hu	1	-----
cDNA_rat	4001	agcatgaattctgtcccaacatctgtatccctggggaaatctcctgtaga
genomic_hu	1	-----
cDNA_rat	4051	caatagtggtcacctgagcatgootgggaccatccaaactgggaaagatt
genomic_hu	1	-----
cDNA_rat	4101	cagtggaaacaacaccacttcccagccccctcagcacaccctcaatacca
genomic_hu	1	-----

cDNA_rat genomic_hu	4151 1	acaagcaca aaattctca agaggaaaactccottgca ccagatctttgt
cDNA_rat genomic_hu	4201 1	aaataaccagaagaaggaggggatgttaaagaatccatatcaattoggtt
cDNA_rat genomic_hu	4251 1	tacaaaagaaccagcgcgaaagotccccaaaatagctcctotttttacco
cDNA_rat genomic_hu	4301 1	acaggtcagagttccccctcagattctacaactctcttgacaagtccggc
cDNA_rat genomic_hu	4351 1	accagctctgtctacaacaatggotgccactcagaacaagggcactgaag
cDNA_rat genomic_hu	4401 1	tagtatoagggtccagaagtctctcagcagggagaagcagcccttcaco
cDNA_rat genomic_hu	4451 1	aa ctcctctcc agtgtcttcttagcaccataagcaagagatctaatacatt
cDNA_rat genomic_hu	4501 1	aaacttctctgtcaacggaaaacccccaaagtgacaagtcctactgtactg
cDNA_rat genomic_hu	4551 1	catctgtcattatgtctgaaacccaacgaacaagatccaaagaagcaaaa
cDNA_rat genomic_hu	4601 1	gaccaaataaaggggcctcggaagaacagaaacaacgcaaacaccacccc
cDNA_rat genomic_hu	4651 1	caggoagggtttctggctatagtgcatactcagctctaacaacagotgata
cDNA_rat genomic_hu	4701 1	cccccttggctttcagtcattccccacgacaagatgatggttggaatgta
cDNA_rat genomic_hu	4751 1	agtgcagttgtttatcactcaacaacotctcttctggccataactgaact
cDNA_rat genomic_hu	4801 1	gtttgagaagtaacaccagacttttgggaatacaacagctttggaacaa
cDNA_rat genomic_hu	4851 1	cgttgtttgagcaaatcacaggagagtacaaacagtgaaaagagcctcagac
cDNA_rat genomic_hu	4901 1	acaccacccaccactcctcagcagtgggggcgccccagtgcccactccttc
cDNA_rat genomic_hu	4951 1	cccacctccttttactaagggtgtgtgttacagacagcaaaagtcacatcag
cDNA_rat genomic_hu	5001 1	ctttccagatgaogtcaaatagagtgtccaccatatatgaatcttcaagg
cDNA_rat genomic_hu	5051 1	cacaatacagatctgcagcaacccctcagcagaggctagccccaatcctga
cDNA_rat genomic_hu	5101 1	gatcataactggaaccactgactctcctctaatctgtttccatccactt
cDNA_rat genomic_hu	5151 1	ctgtgccagcactaagggtagataaaccacagaattctaaatggaagccc
cDNA_rat genomic_hu	5201 1	tctccctggccagaacacaaatatcagctcaagtcatactccgaaaccat

cDNA_rat	5251	tgagaaggggcaaaaggccagcagtaagcatgtccccccacctcagccttc
genomic_hu	1	-----
cDNA_rat	5301	cagaggccagcactcatgcctcacactggaatacacagaagcatgcagaa
genomic_hu	1	-----
cDNA_rat	5351	aagagtgtttttgataagaaaacctgggtcaaaaaccacttccaaacatct
genomic_hu	1	-----
cDNA_rat	5401	gccttacgtctctctacctaagactotattgaaaaagccaagaataattg
genomic_hu	1	-----
cDNA_rat	5451	gaggaaagggtgcaagctttacagttccagctaattcagaogtttttctt
genomic_hu	1	-----
cDNA_rat	5501	ccttgtgagggtgttgagacccactgcccatcatccactggaccagagt
genomic_hu	1	-----
cDNA_rat	5551	ttcatcagganttgaaatatcccaagggaacagaaaagccggttccacg
genomic_hu	1	-----
cDNA_rat	5601	tgcttcccaatggcaccttggtccatccagagggtcagttatcaggaccgt
genomic_hu	1	-----
cDNA_rat	5651	ggacagtacctgtgtctgcatttaatccactgggcgtagaccattttca
genomic_hu	1	-----
cDNA_rat	5701	tgtctctttgtctgtgtgtttttaccocggcaaggattttggacagacatg
genomic_hu	1	-----
cDNA_rat	5751	tcaaggagatcacagttcactttggaagtactgtggaactaaagtgcaga
genomic_hu	1	-----
cDNA_rat	5801	gtggagggtatgccgaggcctacggtttctctggatacttgcaaaccaaac
genomic_hu	1	-----
cDNA_rat	5851	gggtgtctcagaaaacggccaagggaagcagaaagggtotgggtaaacacctg
genomic_hu	1	-----
cDNA_rat	5901	atggascattgatcatotataatctgagtcctttatgatcgtgggtttttac
genomic_hu	1	-----
cDNA_rat	5951	aagtgtgtggccagcaacccatctggccaggattcactgtttggttaagat
genomic_hu	1	-----
cDNA_rat	6001	acaagtcatcacagctccccctgtcattatagagcaaaagaggcaagcca
genomic_hu	1	-----
cDNA_rat	6051	tcgttgggggttttaggtggaagtttgaaaotgocotgcactgcataaagga
genomic_hu	1	-----
cDNA_rat	6101	actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa
genomic_hu	1	-----
cDNA_rat	6151	accattgcagttgactcattccagatttttctgtatccaaatggaactc
genomic_hu	1	-----
cDNA_rat	6201	tgtatataagaagcatcgctccttcagtgggggcaottatgagtgcat
genomic_hu	1	-----
cDNA_rat	6251	gccaccagctcctcaggctcagagagaagggtagtgtattcttactgtgga
genomic_hu	1	-----

rat	7351	CttTGGAAGCAGAGTCAAGGTCCATCCAAATGGAACCTTGGAgatgagga
lc_hu	890	CtaTGGAAGCAGAATCACAGTCCATAAAAATGGAACCTTGGAAattagga
rat	7401	acatccGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGttcggagcgcgag
lc_hu	940	atgtgaGGCTTTCCAGATTTCAGCCGACTTTATCTGTGTGGcccgaaatgaa
rat	7451	ggaGGAGAGAGTGTGTGTGGTAGTGCAGTTAGAAAGTCCTAGAAATGCTGAG
lc_hu	990	ggtGGAGAGAGCGCTGTTGGTAGTACAGTTAGAAAGTACTGAAATGCTGAG
rat	7501	AAGACCAACATTTCAGAAACCCATTCAACGAAAAagtcacgcgcccaagctg
lc_hu	1040	AAGACCGACATTTAGAAATCCATTTAAATGAAAAaatagttgcgcagctgg
rat	7551	gcaagcccgtagCACTGAACTGCTCTGTGGATGGGAACCCCCACCTGAA
lc_hu	1090	gaaagtcacacAGCATTGAATTGCTCTGTTGATGGTAACCCACCACCTGAA
rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCacacaa
lc_hu	1140	ataatCTGGATTTTACCAAATGGCACACGATTTTCCAATGGACCacaaag
rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAAgcaa
lc_hu	1190	ttatcagTATCTGATAGCAAGCAATGGTCTTTTATCATTTCTAAAacaa
rat	7701	ctcggaacaaagtCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGC
lc_hu	1240	ctcgggaggatgCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGC
rat	7751	TAcacgcagaaaactcatcctgttagagATTGGGCAGAAGCCAGTCATTCT
lc_hu	1290	TAtattgcagaaattagtoatattagaaATTGGCCAGAAGCCAGTTATTCT
rat	7801	GACATACGAACCCAGGGatggtgaagagcgtCAGTGGGGAACCGTTATCAC
lc_hu	1340	TACCTATGCACCAGGGAacagtaaaaggoatCAGTGGGAGAATCTCTATCAC
rat	7851	TGCATTGTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
lc_hu	1390	TGCATTGTGTGTCTGATGGaATCCCTAAGCCAAATATCAAATGGACTATG
rat	7901	CCGGGTGGccATGTAATCGACAGGCCTCAAGTGGATGGAAAAATACATAcT
lc_hu	1440	CCAAGTGGtAtGTAGTAGACAGGCCTCAAATTAATGGGAAATACATatT
rat	7951	GCATGAAAATGGCACGCTGGTCAATCAAGCAACAACAGCTcacgcaccaAG
lc_hu	1490	GCATGACAATGGCACCTTAGTCAATTAAGAGCAACAGCttatgcacagAG
rat	8001	GAAATTATATCTGTAGGGCTCAAAACAGTGTGGCCAggcagttattagc
lc_hu	1540	GAACTATATCTGTAAGGCTCAAAATAGTGTGGTCAatacactgattact
rat	8051	gtgtCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAAActacotACC
lc_hu	1590	gttcCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAAAtogtccACC
rat	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
lc_hu	1640	CAGGAGTATTGTCAACCAGGACAGGGGCAGCctttCAGCTCCACTGTGTGG
rat	8151	CCTTGGGAATCCCCAAGCCAaAAGTCACCTGGGAGACGCCAAGCACTCC
lc_hu	1690	CCTTGGGAGTTCCCCAAGCCagAAATCACATGGGAGATGCCTGACCACTCC
rat	8201	CTGCTCTCAAaaagcaacagcaagaaaaccctCATAGAAGTGAGATGCTTCA
lc_hu	1740	CTTCTCTCAacggcaagtaaaagagaggacaCATGGAAGTGAGCAGCTTCA
rat	8251	CCCACAAGGTACGCTgGTCATTTCAGAAATCTCCAAACCTCGGATTCCGGag
lc_hu	1790	CTTACAAGGTACCCtaGTCATTTCAGAAATCCCCAAACCTCCGATTCTGGga
rat	8301	tcTATAAGTGACAGAGCTCAGAACCTACTTGGgaotGATTACGCAACAACT
lc_hu	1840	taTACAAATGCACAGCAAAGAACCCACTTGGtagTGATTATGCAGCAACG
rat	8351	TACATCCAGGTACTCTGACAGGAagggygagactaaaattcaacagaaagt
lc_hu	1890	TATATTCAAGTAATCTGACATGAAataataaagtcaacaacatctggqca

Figure 11

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length	Remarks
1	1	208	208	69	No valid splice site found upstream this exon
2	277	429	153	18	
3	447	485	39	1561	
4	2046	2244	199	1351	
5	3595	3724	130	3254	
6	6978	7359	382	4123	
7	11482	14903	3422	38	
8	14941	15307	367	51	
9	15358	15825	468	1039	
10	16864	~17760	~897	?	Last exon might be complete
11	?(1)	2317	~231 7	25	Exon is not complete and start site is not known
12	2342	2397	56	-	

Figure 12

ATGAAGGTAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC
CCTGGGGGCAAGGCCTGTCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTTCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC
ATTACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACCTTCAGAAAGATACCTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
AACAAATATTGAGTTTATAAACCCAGAGGTTTTTATGGGCTCAACTTTCTCCGCCTGGTGCCTTGGAA
GGAAATCAGCTCACTAAGCTCCACCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC
TCTTTCATTAAGTTCCTATACTTGTCTGATAACTTCTTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGGATACAGCCAGATGTAATAAAATGCAAAAAAGATAGAAGTCCCTCTAGTGCCTCAGCAGTGT
CCACTTTGCATGAACCCCTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAG
TGTGCCAAGCCAACCATTTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCT
GCTTTTCATCTCTCCCAAGGTTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGATCAGTCT
GGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCCATTGCATTCACTGAA
GAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAGATTACGGTCAC
ATTGAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCTCTGATACTAGAAAGGAGCCACTTG
CTTAGTGAAACACCCGAGCTCTATTACAAATATAAACAGGTGGTTCCTAAGCCTGAAGACATTTTACC
AACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAGCTGAAC
AGAATGCCACCACATTGAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGA
GCAGAGATGAGGCCAGTGAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACAT
ACTGCTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCACCCACACGTTGGAT
TGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCCTTATGTCAGTGAGGATGGACGGATCCTAATAGAC
AAAAGTGGAAAATTGGAACCTCAGATGGCTGATAGTTTTGACACAGGCGTATATCACTGTATAAGCAGC
AATTATGATGATGCAGATATTCTCACCTATAGGATAAAGTGTGTAGAACCTTTGGTGAAGCCTATCAG
GAAAATGGGATTTCATCACACAGTTTTTCATTGGTGAACACTTGATCTTCCATGCCATTCTACTGGTATC
CCAGATGCCTCTATTAGCTGGGTATTTCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAA
GTTCTAAACAATGGCACATTAAGAATATTACAGGTCAACCCGAAAGACCAAGTTATTATCGTGTGTG
GCAGCCAAACCATCAGGGGTTGATTTTTTGTATTTTCAAGTTTCAGTCAAGATGAAAGGACAAAGGCC
TTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTTAAGGAG
CCACCAGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACA
AGTAAGAGGCACAACATATCGGGAATTAACTACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGG
GAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACGTGTGGAG
AAAGCTAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCACCCCAAGTGGTCAAC
CAACTCCCAAACATACCTGGTGAAGAAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATTTATG
GTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAACAATATCTGAT
AGTCTATGACAAACATAAATTATGGCACAGAATTCTCTCTGTTGTGAATTCAAAATACTACCACCT
GAAGAACCCACAGATTTCAAACCTGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAATATAAACCCA
ACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCACTGCTCTTTCCACTGCTACTTTGGA
GCAACTGAATTTGAGGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCCCAATA
ACAGTAAGGACTATGATCAAAGATGTCAATGTCAAATGCTTAGTAGCACCACCAACAACTATTATTA
GAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCAC
TTCTATTCTCACAATACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACAGCTGCTCAT
TCTCAGTTTTCCGATCCCTAGAAATAGTACAGTTAAACATCCCGCTGTTTCAGACGCTTTGGGAGGCAGAGG
AAAATTGGCGGAAGGGGGCGGATTATCAGCCATATAGAATCCAGTTCTGCGACGGCATAGATACAGC
ATTTTCAGGTCAACAACAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAAT
GTGACATGCTGTCTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACCTTAGTC
CAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAAACAACCCCAATAAAATATTTT
AGGACTGAAATTTCCCAAGTGACTCCAACCTGGTGCAGTCATGACATATGCTCCAACATCCATACCCATG
GAAAAAATCACAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACAATGAAGCTAAAAGAGATTCA
GTGATTACATCGTCACTTTTCAAGTGCTATCACCAGCCCAATGACTATTATAGCCATTACAAGGTTT
TCAAGAAGGAAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAAGGAAT
CAACATAAAGTTAGTTTACAAAAAAGCAGCTGTGATGCTTCTTAAACATCTCTGCTTTTACCACAG
AGACAAAGTTCCCTTTCCATTTTACCACACTTTTCAACAAGTGTGATGCAAAATTCATCTAATACCTTG
ACTACCGCTCACCACACTACGACCAAAACACACAATCCTGGAAGTCTTCAACAAAGAAGGAGCTTCCC
TTCCCACCCCTTAACCTATGCTTCTTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATCATATCA
ACGCAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCACTTATGAAACCCAAACA
GAGAGATCTAGAGCACAACAATACAAAGAGAACAGGAGCCTCAAAGAAGAACAGGACTGACCCAAAC
ATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACT

CATTCCCCACCAGAAAACACAACCTGGGATTTCAAGCACAACTCAGTTTTTCATTCAAGAACTCTTAATCTG
ACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCTGAAACA
ACTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCATTAGACACTCCCATACCACCA
TTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCAGTTACT
GACACACGTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGG
CACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT
ACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAATCCATCTCCCTGGGCAGAA
TACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCCAAAAAGCCAGAAGTAAGCATGTTG
GCTACTACAGGCCTGTCCGAGGCCACCACTCTTGTTCAGATTGGGATGGACAGAAGAACAAGAAG
AGTGACTTTGATAAGAAACCAGTTCAGAAGCAACAACCTCCAACTCCTTCCCTTTGACTCTTTGTCT
AGGTATATATTTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTACTATTCCAGCTAATCA
GATGCCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCACCAATTCATTGGACCAGAGTTTCAGGA
CTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCAATGGTACCCTGTCCATCCAG
AGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTT
CATGTACACCTTGTCTGTGGTTTTCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTT
CATTCCCGAAGCACTGTGGAAGTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACCTGGATT
CTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGA
ACATTGGTCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAACCCAGGT
GGCCAGGATTCAGTGTGGTTAAAATACAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAAAGGAGG
CAAGTCATTGTAGGCACCTTGGGGTGAAGTTTTAAACTGCCCTGTACTGCAAAAGGAACCTCCTCAGCCC
AGCGTTTACTGGGTCTCTCTGATGGCACTGAAGTGAACCATACAGTTTACCAATTCCAAGTTGTTT
TTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACTTATGAATGCATT
GCTACCAGTTCCACTGGTTCCGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCC
AGGATAGAAGCTGCATCCCAGAAAAGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAAGTGTCTCA
GCCACTGGGGAGCCCAAACCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTTCGACCAGTGGAGC
TGGATCCACGTCTACCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTCTAC
TTGTGTGTGGCAAGAAACAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTGAAACCT
GCCAAAATTGACCACAAGCAGTATTTTAGAAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGC
AAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCTGATGGAACCATGATCAACAATGCA
ATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCTTTTCAACAATGGAACCTTTATACCTC
AACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACCTTGCTATGCCCAGAACACCCTAGGGAAAGATGAA
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGATAAGGCAGAGTAACAAAACCAACAAGAGA
ATCAAAGCTGGAGACACAGCTGTCTTGAAGTGTGAGGTCACTGGGGATCCCAAACCAAAATATTTTGG
TTGTGCTCTTCCAATGACATGATTTCTTCTCCATTGATAGGTACACATTTTATGCCAATGGGTCTTTG
ACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCAGTGGGGAT
GACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACAAACAGA
ACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCTGAAGGGACACCA
TCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTTCTCAGACCCCATACTATGGAAGCAGAATC
ACAGTCCATAAAATGGAACCTTGGAAATTAGGAATGTGAGGCTTTTCAAGTTTCAAGCCACTTTATCTGT
GTGGCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGA
CCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTGAATTGC
TCTGTTGATGGTAACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACACGATTTTCCAATGGA
CCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTCTTAAACAACTCGGGAGGAT
GCAGGAAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATT
GGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAGAATCTCTATCA
CTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTATGTAGTA
GACAGGCCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAAGCAACA
GCTTATGACAGAGGAACTATATCTGTAAGGCTCAAAATAGTGTGGTGCATACACTGATTACTGTTCCA
GTAATGATTGTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGGACAGGG
GCAGCCTTTTCACTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACATGGGAGATGCCTGAC
CACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGAAGTGAGCAGCTTCACTTACAAGGTACC
CTAGTCATTTCAGAATCCCCAAACCTCCGATTCTGGGATATACAAATGCACAGCAAAGAACCACCTTGGT
AGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

Figure 14

rat_cDNA	CGAGAGACGACAGAGGTTACGGCTGCGGAGAAACGACGAGAGGGTCCAGAAAAAGAAA
human 5+3 corrected	-----CAGAAGGGTCCAGGAAA-GGAAA
mus_cDNA_5	
rat_cDNA	GTGCTGGAGGGGAGTGGGACAAAAGCAGCGACCAAGTGAATGTCACTTCAGTGACTGAG
human 5+3 corrected	-----GTACTGGAGGGGAGTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAA
mus_cDNA_5	
rat_cDNA	GCCAGGCAAAAGCGCGGGAAGGATTTTGTGTAGCTTGGGACCTTTTCATAGACTGAT
human 5+3 corrected	-----GCCAGGCAAAAGGAGCGGGAAGGATTTATATGTAGCTGGGACGCTTTTCATAAACTGAT
mus_cDNA_5	
rat_cDNA	GACACGTTTACGCAAAATA-GAAATTTGAGGAGAAACGCTGGGCTTCGGAAAG---GA
human 5+3 corrected	-----GACGTGTTTGTGCAAAAGCAAGCAATTTGAGGAGAAACGCTGGGACGTTCGGAAAGAAGGA
mus_cDNA_5	
rat_cDNA	GTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAACTAACTAATGTATACTA
human 5+3 corrected	-----GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTT---GAGAACTAACTAACCTATACTA
mus_cDNA_5	
rat_cDNA	TTGAGGGAGGAGGAAGAGCATTACAGAGTTTCAGCAGCAGCAGGAAGCTTTGTTAAT
human 5+3 corrected	-----TTGAGGGAGGAAGAGAGCATT-----CCAGCAGCAGCAGGAAGCTTTGTTAAT
mus_cDNA_5	
rat_cDNA	TTGGAATGGATGATAGCATTAAATAACAGAAGCGCTCCAGGTCTCTGAAGCTTCAGT
human 5+3 corrected	-----TTGGAATGTATGATACCATTAATAACAGAAGCGCTCCAGTTCTCTGAAGAGTCAGT
mus_cDNA_5	
rat_cDNA	CCCCCAGCTGAAAGCCAGAAAAGACTAAGCCCACTAAGCCTTTTGTATCCCTTTGGAAGCA
human 5+3 corrected	-----CCCCCAGCT-----A-GTGTAAGCCTACTAAGCCTTTTGTATCCCTTTGGAAGCA
mus_cDNA_5	
rat_cDNA	AAGAAGTTTCCTTCCCTGGGGTGAAGACTCTCTCAGAGATTTCCTGTCTCTGCCTATG
human 5+3 corrected	-----AAGAAGCTTCTTCAATCAGGTGAAGGCTCTCTCAGAGATTTCCTGTCTCTGCCTATG
mus_cDNA_5	
rat_cDNA	TTACAAGAGGAATCAAAACAGACAGAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAG
human 5+3 corrected	-----ATGAAGGTAAAGGCAGAGGAA
mus_cDNA_5	TTACAAGAGGATTCAAAAGCAAGACAGAGAGCTCAGGATGCAGAGAGAGGCAGGGAAG *** ** * ***** *
rat_cDNA	TCAGCGGCTTGTTGATCTCCCTCACTGCTGTCTGCTGGTGGTCACTCCCTGGGAGCAGGG
human 5+3 corrected	-----TCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCTGGTGGTCACTCCCTGGGAGCAGGG
mus_cDNA_5	TCAGCTGCTTGCTGATCTCCTCACTGCCATCTGCTGGTGGTCACTCCCTGGGAGCAGGG *** * ***** * ***** * ***** * ***** * ***** *
rat_cDNA	CCTGTCTCGCGGCTGTGCTGCTATGTGCCCCAGAGGTGCCTGTACATTTGGGTACC
human 5+3 corrected	-----CCTGTCTCGCGGCTGTGCTGCTATGTGCCCCAGAGGTGCCTGTACATTTGGGTACC
mus_cDNA_5	TCGTCTCTCGCGGATGTGCTGCTATGTGCCCCAGAGGTGCCTGTACATTTGGGTACC ***** ***** * ***** * ***** * ***** * ***** *
rat_cDNA	TGACCTCCATCCAGATGG-CATCCCGCCCATGTGGAAAGATTAATTTAGGATATAAC
human 5+3 corrected	-----TGACTTCCATCCAGACAG-CATCCCGCCCATGTGGAAAGCATCAATTTAGGATATAAC
mus_cDNA_5	TGACCTCCATCCAGAGCGGCTATCCAGCCATGTGGAAAGATCAATTTAGGATATAAC *** ***** * ***** ***** * ***** * ***** * ***** *

00946370-030466

AGCCTTACTAGATTGACAGAAAAAGACTTTGATGGCCTGAGCAAACTGGAGTTACTCATG
AGCTTCGTTAGATTGATGGAAACAGATTTTCTGCGCTGACCAAACTGGAGTTACTCATG
AGCCTCAGTATGACAGAAAAATGACTTTCTGCGCTGAGCAGACTGGAGTTACTCATG
*** * ***** ** * ***** * ***** *

CTGCACAGTAATGGCATTTCACAGAGTCAAGAACCTTCTCGGGCTTGCAAGTCCCTTG
CTTCACAGCAATGGCATTTCACAAATCCGTGACAGAACCCTTCTCAGGCTTGCAAGGCCCTTG
CTGCACAGCAATGGCATTTCACAGAGTCAAGAACCTTCTCGGGCTTGCAAGTCCCTTG
** ***** * ** ***** *

CAGGTCCTTAAAAATGAGCTATAACAAAGTCCAAATCATTCCGAGGATACTTTCTACGGA
CAGGTCCTTAAAAATGAGCTATAATTAAGTCCGAAAACCTTCAGAAGATACTTTTATGGC
CAGGTCCTTAAAAATGAGCTATAACAAAGTCCAAATTAATTCGAGGATACTTTGTATGGA
***** ** **

CTCGGGAGCTTGSTCGGTTGCACTGTGGATCAACAACAAATTGAATTGATCAACCTTGAG
CTCAGGAGCTTGACACGATTTGCACATGGACACACAAATATGAGTTTATAAACCAGAG
CTCAGGAGCTTGACCCGTTGCACTGTGGATCAACAACATTGAGTTTATCAACCCGAG
*** ***** ** ***** ***** ***** ** ***** **

GCGTTTATGGACTACCTCGCTCCGCTTGGTACATTAGAGGAACCGGCTCACAAAG
GTTTTTATGGGCTCAACTTTCCTCGCCTGGTGCATCTGGAAGGAATCAGCTCACTAAG
CGGTTTACGGCATCACCTTGCCTCGCTTGGTACATCTAGAGGAACCGGCTCACAAAG
* * * * * * * * * * * * * * * * * * *

CTCCATCCGAGACACATTGTCTCATTAAAGCTATCTCCAGATATTTAAACCTCTTTTCATT
CTCCACCCGAGATACATTGTCTCTTTGAGCTAAGCTCCAGATATTTAAATCTCTTTTCATT
CTCCATCCGAGACACATTGTCTCTTTGAGCTATCTCCAGATATTTAAACCTCTCTTCATT

AAGTACCTGTTCTGTCTGATAACTTCTCT-GACCTCCCTCCCAAAGAAATGGTCTCCTA
AAGTTCCTATACTGTCTGATAACTTCTCT-GACCTCCCTCCCTCAAGAGATGGTCTCCTA
AAGNACCTGTACTGTATGATAACTTCATGACCTCCCTCCCAAAGAAATGGTCTCCTC
*** 444 4 ***** 4***** 4***** 4***** 4***** 4*****

CATGCGCAACCTAGAAAGCCTGTATTGCAATGGAAACCCATGGACCTGTGACTGCCATT
 TATGCGCTGAOCTAGACAGCCCTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATT
 TATGCGCAACCTAGAAAGCCTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATT

AAAGTGGTGTGCTGAGTGGATGCAGGGAACCCAGATATATAAAATGCRAGAAAGACAG
AAAGTGGTGTGCTGACTGGATACAGGNNNNNCCAGATGTATAAAATGCAAAAAGGTAG
AAAGTGGTGTGCTCGAGTGGATGCAGGGAACCCAGGTA-ACATCTTGT-----TTGTTTG
***** ** ***** ** * * * * *

AAGTCTTCCAGTCTCTCAGCAATGTCCCT-TTGCATGAACCCGAGATCTCTAAGGCA
AAGTCCCTCTAGTCTCTCAGCAATGTCCCT-TTGCATGAACCTAGGACTCTAAGGCA
TTTCITTTTTTATARKAGTATTTCTCAATTTCAATTAGATGATATCCCAAGTCTC-
* * * * * * * * * *

GACCCCTTGCTAGGTACCATCTCGAGCTTTCCTATGTACAAAGCCACCACTTGATCCAT
ACCGCTTAGCTATGGTCTCAGCTCGAGCTTTCAGTGTGCCAAGCCACCACTTGACTCAT
-CCCATACCTCCCCCA-----CTTCCCTACCTACCACTTC--CCATTTTTCG

 ** * * ** * * * * *

CACTGAAGTCAAAGAGCCTGGTACTACGGAGGCAATGGATCTGCGCTCCACCTCACTC
CCCTGAATCAAAGAGCCTGACTATTCTGGAGACAGTAGTCTGCTTTCACTCTCCCC
CCTGGCATTCCCC-----
* * * *

AAGATTTCATAGAACCCCTTGGCTCCTTGCTCTTGAACATGACANANNTNTCTGGAATA
AAGGTTTCATGGCACCCCTTGGCTCCTCACTTTGAATATGACAGATCGTCTGGAATG

AGGCCGACATGGTCTGTAGTATCCAAAAGCCATCAAGGACATCACCAACTGCATTCACTG
AAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCCATTGCATTCACTG

AGGCCGACATGGTCTGTAGTATCCAAAAGCCATCAAGGACATCACCAACTGCATTCACTG
AAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCCATTGCATTCACTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAGAAATGACTACATCATGCTAAATGGTCATTTTCCACAAATCTTGTGTCCAGTGTAG
AAGAAATGACTACATCGTGCTAAATACCTTCATTTTCAACATTTTGGTGTGCAACATAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATTATAATCACATCCAGCCAGTGTGGCAACTTCTGGCTTTATACAGTGACTCTOCTCTGA
ATTACGGTCACATTCAGCCAGTGTGGCAAAATTTGGCTTTGTACAGTGATTCTOCTCTGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TACTAGAAAGGAAGCCCCAGCTTACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGG
TACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGCTCTTAGGCTGAAGACATTTTACACGATAGAGGCTGATGTCTAGAGCAGACCOCTT
TGGCTCCTAAGCTGAAGACATTTTACCAACATAGAGGCGAGATCTCAGAGCAGATCCCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTGGTTCCACACAGAAAAAATGTCTTGCACTGAACAGAACTGCCACACACTTAGCA
CTTGGTTAATGCAAGACCAAAATTTCTTGCACTGAACAGAACTGCCACACACTTAGTA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CATTACAGATCCAGTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAG
CATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGAGCAGAGATGAGGC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CGGAGAGACTCAAATGGACCATGATCTGATGATGAACAATCCCAAACCTGGAAACGCACTG
CAGTGAAACACAAATGGACTATGATTTCAAGGGATACAAATACTAAGCTGGAAACATACTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TCTTGGTTGGCGGCACTATTGCOCTGAGCTGTCCAGGCAAGGGGACCOCTTCAOCTCACT
TCTTGGTAGGTGGAAACCGTTGGCTGAACTGCCAGGCGAAGGAGACCCACCCACACG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCCTTACGTTAGGAGGATGGGC
TGGATTGGCTTCTAGCTGATGGAGTAAAGTGAGAGCCCCCTTATGTCACTGAGGATGGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGATGCAG
GGATCCTAATAGACAAAAGTGGAAAAATGGAACTCCAGATGGCTGATAGTTTTCACACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GTCTTTACCACTGCATAAGCAACCAATGATGCAGATGCGGATGTTCTCACATACAGGATAA
GCGTATATCACTGTATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTGGTAGAGCCCTATGAGAGAAAGCACACATGACAGTGGAGTCCAGCACACAGTGGTTA
CTGTGGTAGAACCTTTGGTGGAGCCCTATCAGGAAATGGGATTCATCAGCAGTTTCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CGGGTGAGAGCGCTGACCTTCCATGCTTTCCACGGGTGTTCCAGATGCTTCTATTAGCT
TTGGTGAAACACTTGATCTTCCATGCAATTTCTACTGGTATCCAGATGCCCTATTAGCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGATTCTTCCAGGGAACTGTGTCTCTCAGCCATCAAGAGACAGGCAATTTCTAACA
GGGTTATTCCAGGAAACATGTGCTCTATCAGTCATCAAGAGACAGAAAGTTCTAACA

09501630-030402

rat_cDNA human 5+3 corrected mus_cDNA_5	ATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGATCAAGGTCATTACCAATGTGTGG ATGGCACATTAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTGG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCCAACCCATCAGGGGCGGACTTTTCCAGTTTAAAGTTTCAGTTCAAAGAAAGGCG CAGCCAACCCATCAGGGGTTGATTTTGTATTTTCCAGTTTCAGTCAAGATGAAGGAC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AAAGGATGGTTGAGCATGACAGGGAGGCAGGTGGATCTGGACITGGAGAACCCAACTCCA AAAGGCCCTTGGAGCATGATGGAGAACAGAGGGATCTGGACTTGATGAGTCCAACTCTA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGTTTCCTTAAGCAGCCAGCATCTTTGAAACTCTCTGCATCAGCTTTGACAGGCTCAG TTGCTCATCTTAAGGAGCCACAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AGGCTGGAAAACAAGTCTCCGGTGTACATAGGAAGAACAAACATAGAGACTTAATACATC AGGTTGGAAAACACACTCAAGCACAGTAAGAGGCACAACTATCGGGAAATTAACACTCC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GGCGGCGTGGGGCAITCCACGCTCCGGGCAITCAGGGAGCATAGGAGGCAGCTCCCTCTCT AGCGACGTGGAGATTCAACACATCGACGTTTTAGGGACAATAGGAGGCATTTCCTCCCT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCTCGGAGAAATGACCCGCAACGCTGGGCAGCACTTCTAGAAAAGCCAAAAGAAAT CTGCTAGGAGAAATGACCCACAACTTGGGCGGCACTGTTGGAGAAAGCTAAAAGAAATG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTGCCAAAAAGCAAGAAATACCAAGTAAGCCAGTGCCACTGGCTGTTCCCTCG CTATGCCAGACAGCGAGAAATACCAAGTGAGCCACCCCAAGTGGTCACCCAACTCC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGAACTCACTGACGAGGAAAGGATGCTCTGGCATGATTCCTCCAGATGAAGAATTCA CAACATACCTGGTGAAGAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATTGA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGTTCTGAAACTAAGGCTTCIGGTCTCCAGGAAGGTCAACCACTGCTGACTCTGGAC TGGTCCCGGCCACTAAAGCTTTGAACCTTCAGCAAGGACAGTGACTGCTGACTCCAGAA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGTAAATCATGGTTTTATGAAGATATAGCTTCTGGCACAGAGTCTCAACTGTGAATC CAATATCTGATAGTCTATGACAAACATAAATTATGGCACAGAACTCTCCGTTGTGAATT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CACAAACACTACATCTGAGCACCTTCTGATTTCAAAATATTTAGTGTAAACAAAGGTA CACAAATACTACCACTGAGAAACCAAGATTCAAACGTCTACTGCTATTAAACTA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACAAACACC CAGCCATGTCAAAGAATATAAACCCAAACATGTCAAGCCAAATACAAGGCACAACCAATC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AAAACCCATCATTAATCTTTCCATC-----AGTAGCTGAATTCGAGATTCTG-CT AATATTATCACTGTCTTTCACTGCTACTTGGAGCACTGAATTTAGGACTCTGACA -----

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGCA---GGAAGAGCATCTTCCCAAAGTGCAACCCCTGTACAGGGGGAAACATGGCT
GAGGGGAGAGGAAGAGAGCATTTCC--AGTA-AOCCCCAATAACAGTAAGGACTATGATC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACCTATGGCCATACCAACACATA--TAGTAGCTTTACCAGCAAAGCCAGTACAGTCTTTC
AAAGATGNTCAATGTCAANATGCTTAGTAGCAOCCCAACAA-CTATTA-----TTAG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGCCATATATCCACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCAGCAGAC
AGTCAGTAATACCAAAATAGTCAT-----CAGACATCTGTAGAGAGTGAAGTGAAC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CTAGCAGTAGTGACATCTCTTCTCAGCTACTGCAGACCCCTAGCTTCTCCAGTCACCCCT
CCAGGCACATCACTTCTATTCTCAGCTACTCAAATACCTTAGCACCTCCAGCTTCCCTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGTTACACACACCACTGCTCTCTTTATTTACATTCCTAGAAACAAATACAGGTA
CAGATCCACACACAGCTGCTCATTTCTCAGTTTCGATCCCTAGANNNAATAGTACAGTTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACTTCCCTTGTCCAGGCCTTGGGAAGAGAGAGGACAATTGGAGCAGAGGGAGAGTTA
ACATCCCGCTGTTCAGACGCTTGGGAGGCAGAGGAAATTTGGCGGAAGGGGGGGATT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAAACCCACATAGAACCCAGTTCTCCGACGGCATAGACAGGACTGTGAGGCCAGCAA
TCAGCCCATATAGAACTCCAGTTCTGCGAAGGCATAGATACAGCATTTCAGTCAACAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCAGGGACCTGCTAACAAAATGTGAGCCAGTTCCAGCCACAGAGTACCCCTGGGATGT
CCAGAGGTTCTCTGAAAAAGCACTACTGCACTTCTCAGCCACAGTGTCAATGTGACAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCCACACATGTCTTCCGAGAGGGGCTCAGTGGCTACTGCAGCACTGTCASTTCCAA
GTCTGTCTGTCTTCCAGGGAGAGGCTCAOCCACTGCCACAGCAGCATTGTCTTTCCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTTTCATCCACAGTGGCTCCCAAACTAAATATGTGGGGTCTATGCAGAGAGTCTA
GTGCTGTCTCCATCACCTTCCCAAGCTGACATTGCTAGAGTCCCATCAGAGAGTCTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CCACTGTGGTCAAGAAACCACTGTACTATTTAAGGACAAACAAATGTAGATATTGAGA
CAACTCTAGTCCAGATCCACTATTACTACTTGAGAACAAACCCAGTGTAGANNNGAAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TAATACAAOCCACTAATAAATATTCGGAGGGGAAGTAACACGTGATTCCTACGGAG
NNACACACCCACAATAAATATTCAGGACTNGAAATTTCCAAAGTGAATCCAACTGGTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGCATGACTTCTGCTCCACATCTGTATCCCTGGGGAAATCTCTGTAGACAATAGTG
CAGTCATGACATATGCTCCACATCCATACCCATGGAAAAAAGTCAAAAGTAAAGGCA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTCACCTGAGCATGCTGGGACCATCCAACTGGGAAGATTCACTGGAACACACACCAC
GTTACCCACGTGTGTCTAGACCAATGAAGCTAAAGAGATTCACTGATTACATCGTCAC

0099130-030402

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTCCAGCCCCCTCAGCACCCCTCAATACCA-----CAAGCACAAATTCCTCAAGA
TTTCAGGTGCTATCACCAAGCCCAATGACTATTATAGCCATTACAAGGTTTCAAGAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGAAACTCCCTTGCAACAGATCTTTGTAATAACCAAGAAGSAGGGGATGTTAAAGA
GGAAATTCCTGGCAACAGACTTTGTAATAACCAAGCAAGGCGATTAAAGGA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATCCATATCAATTGGTTTACAAAAGAACCCAGCCGCAAGGTTCCCAAAATAGCTCCTC
ATCAACATAAAGTTAGTTTACAAAAAGCACAGCTGTGATGCTTCTCTAAACATCTCCTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTTACCCACAG-GTCAGAGTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCGGOCA
CTTTACC-ACAGAGACAAAGTCCCTTTCCATTTCACCACACTTTCAACAAGTGTGATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCAGCTCTGTCTACAACAATGGCTGCCACTCAGAACAGGGCACTGAAGTAGTATCAGGT
CAATTCCTATCTAATACTTGACTACCGCTCACCACTACGACAAA--ACACACAA-T

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCCAGAAGTCTCTCAGCAGGGGAAGAGCAGCCCTTCAOAACTCTCTCCAG-TGCTTCC
CCTGGAAGTCTTCCAAAGAGAGGAGCTTCCCTTC-CCACCCCTTACCCCTATGCTTCC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TAGCACATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCACAGT
TAGTATTATAAGCAAGACTCAAGTACAAAAGCATCATATCAACGCAACAGCAACCGC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GACAAGTCCTACTGCTACTGCATCTGTCTATTGTCTGAAACCAACGAACAGATCCAA
AACAACTCCTACCTTCCCTGCATCTGTCTACTTATGAAACCAACAGAGAGATCTAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGARGCAAAGACCAATTAAGGGGCTCG--GAAGAACAAGCAAGCAACACCCAC
AGCACAAACAATACAAAGAGAAGGAOCTCAAAGAGAAGCAGGACTGACCCAAACATCTC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCCCAGGCAGGTTTCTGGCTATAGTGCACTCAGCTCTAACACAGCTGATAACCCCTT
TCCAGACCAAGATTCTGGCTTCACTACCCCACTGCTATGACNACCTCCTNNGCTCTNN

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGCTTTCAGTCATTCCCCACGACAGATGATGGTGGAAATGTAAGTGCAGTTGCTTATCA
NGCATTCACTCAITCCCCACAGAAAACCAACTGGGATTTCAAGCACAATCAGTTTTC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCAACAACCTCTCTTCTGGCCA--TAACTGAAGTGTGAG-AGTAC--AACCAGAC
TTCAAGAACTCTTAATCTGACAGATGTGATTGAGAAGTACGCCAAGCAAGTACTCAGAC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTTGGGAATACAAAGCTTTGGAAACAAAGTTGTGAGCAATCAGAGAGATACCAC
TTTGAAGAGCACATTTGCTTCTGAAACAACTTTGTCCAGCAATCACCAGAGTACCAC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGTGAAGAGAGCTCA--GACAC--ACCAACACACTCCTCAGCAGTGGGGGGCCCCC
AACTAGGAAGCATCATTAGACACTCAACCAACCACTTCTTGAGCAGCAGTGTACTCT

CGCTGCTGGTGG

rat_cDNA human 5+3 corrected mus_cDNA_5	AGTGCCTACTCTCTCCACCTCTCTTTACTAAGGGTGTGGTTACAGACAGCAAGTCAC AATGCCAGTTCCCATCTCTCCCTCCCTTTACTCAGACAGCAGTTACTGACAACGTGGGAC
rat_cDNA human 5+3 corrected mus_cDNA_5	ATCAGCTTTCCAGATGACGTCAAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA TCCCATTTCCGGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAA
rat_cDNA human 5+3 corrected mus_cDNA_5	TACAGATCTGCAGCAACCTCAGCAGAGGCTAGCCCCAATCCTGAGATCATAACTGGAAC TCCNNNNNNNCA--AATGCCAAGTCA-----CNAATTTGNGAACNNNNACTCNNNN
rat_cDNA human 5+3 corrected mus_cDNA_5	CACTGACTCTCCCTCTAATCTGTTTCCATCCACTTCTGTGCCAGCACTAAGGTTAGATAA NACTTCATCTACNTCTAATCTGTACATTCTACTCCCATGCCAGCACTAACACAGTTAA
rat_cDNA human 5+3 corrected mus_cDNA_5	AACCACAGAAATCTAAATGGAAGCCCTCTCCCTGGCCAGACACAAATATCAGCTCAAGTC ATCAGCAGAAATCCAAATTAAGTCCATCTCCCTGGGCAGAAATACCAATTTTGGCACAAGCC
rat_cDNA human 5+3 corrected mus_cDNA_5	ATACTCCGAAACCAATTGAGRAGGGCAAAAGGCCAGCAGTAAGCATGTCCCCCACCCTCAG ATACTCAGACATTGCTGAAAAGGCCAAAAGCCAGAACTAAGCATGTTGGCT-ACTACAG
rat_cDNA human 5+3 corrected mus_cDNA_5	-CCTTCCAGAGGCCAGCACTCATGCTCACACTGGAAATACACAGAAGCATGCAGAAAAGA GCCTGTCCGAGGCCACCACTCTTGTTCAGATTGGGATGGACAGAAGAACACAAAGAGA
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGTTTTTGAATAAGAACTGGTCAA-AAAC--CAACTTCCAAACATCTGCCTTACGCTCT GTGACTTTGATAAGAAACCACTTCAAGAGCAACAACCTTCCAAACTCCTTCCCTTTGACT
rat_cDNA human 5+3 corrected mus_cDNA_5	CTCTACCTAAGACTCTATTGAAAAAGCCAGAAATAATTGGAGGAAAGGCTGCAGCTTTA CTTTGTCTAGGTATATATTGAAAAGCCAGGATAGTTGGAGGAAAGCTGCAAGTTTAA
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGTTCAGCTAATTCAAGCTTTTCTTCTTGTGAGGCTGTGGAGAACCACTGCCCA CTATTCCAGCTAACTCAGATGCTTCTTCTCTGTGAAGCTGTGGAAATCCCTGCCCA
rat_cDNA human 5+3 corrected mus_cDNA_5	TCATCCACTGGACCAGAGTTTCATCAGGANITGAATATCCAGAGGACACAGAAAGCC CCATTCATTGGACAGAGTNNNNTCAGGACTTGATTATCTAAGAGGAAACAGAAATAGCA
rat_cDNA human 5+3 corrected mus_cDNA_5	GGTCCACGTGCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGAACGTG GGGTCCAGGTTCTCCCAATGGTACCCCTGTCCATCCAGAGGGTGGAAATTCAGGACCGCG
rat_cDNA human 5+3 corrected mus_cDNA_5	GACAGTACCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTGT GACAGTACTTGTGTTCCGCATCCCAATCTGTTTGGCAGAGCCACCTTCATGTCACTTGT
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTGGTTTTTTTACCCGGCAGGATTTTGGACAGACATGTCAAGGAGATCAGGTTCACT CTGTGGTTTCTCTATCTCCAGGATCTCTGAGAGAGTACCAAGAGATCAGGTTCACT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTGGAACTACTGTGGAACTAAAGTCAGAGTGGAGGGTATGCCGAGGCTACGGTTTCCT
CCGGAAGCACTGTGGAACTGAAGTCAGAGCGAGAGGTAGGCCAAGCCCTACAGTTACCT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGTACTTGCACCAACCAACGGTGGTCTCAGAAACGGCCAAAGGGAAGCAGAAAGGTCTGGG
GGATTCTTGCAAAACCAACAGTTGTCTCAGAAATCATCCAGGGGAAGTAGGCAGGCTGTGG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TAACACCTGATGGAACTTGCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA
TGACGGTTGACGGAACTTGGTCTCCCAATCTCAGTATTTATGACCGTGGCTTTTACA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACTGTGGTTAAGATACAGTCATCA
AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGTGGTTAAATACAGTCATTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGCTCCCCCTGTCAATTATAGAGCAAAAGAGGCAAGCCATCGTTGGGGTTTTAGGTGGAA
CAGCACCACCTGTTATTCTAGAGCAAAAGGAGGCAGTCATTGTAGGCACCTGGGGTGAAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GTTTGAAACTGCCCTGCACTGCAAAAGGAATCCCCAGCCTAGTGTTCAGTGGGTCTTT
GTTTAAACTGCCCTGTACTGCAAAAGGAATCCTCAGCCAGCGTTTACTGGGTCTCT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATGATGGGACTGAACATAAAACCATTCAGTTGACTCAATCCAGATTTTCTGTATCCAA
CTGATGGCACTGAAGTGAACCATTCAGTTTACCAATTCAGTTGTTCTTAMTTTCAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATGGAATCTGTATATAAGAAGCATCGCTCCTTCAGTGAGGGGCACCTATGAGTGCATTG
ATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACCTATGAATGCATTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCACCGCTCCTCAGGCTCAGAGAGAGGGTAGTGATTCTTACTGTGGAAGAGGGAGAGA
CTACAGTTCCACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAAAGGAAGAGCGAGTGA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAATCCCAGGATAGAACTGCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT
CCAGCCCAGGATAGAACTGCTCCAGAAAGGACTGAAGTGAATTTGGGGACAAAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TACTACTGAACTGCTCAGCTACTGGGGATCCAAAGCCTAGAATAATCTGGAGGCTGCCAT
TACTACTGAACTGCTCAGCCACTGGGGAGGCCAAACCCCAATAATGTGGAGGTTACCAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCAGGCTGTCTACGACCACTGGCAAGAAATGGGCGCCGAATCCAGCTACCCAAATG
CCAGGCTGTGGTCAACCACTGG-----GCAGCTGGATCCAGCTACCCCAATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GATCCTTGGTGGTTGGGTCACTGACGGAAAAAGACGGTGGTCACTACTTATGTGTGGCAA
GATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTCTACTTGTGTGTGGCAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GAAACAAAATGGGAGATGACCTAGTCTGTATGATGTCCGCTGAGATTGACACCTGCCA
GAAACAAAATGGGGATGATCTGATCTGATGATGTAGCCTAGACTGAAACCTGCCA

05994630-030402

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAATTGAACAGAAGCAGTATTTTAAGAAGCAAGTGTCCATGGGAAAGATTTCAGTTG
AAATTGACCAACAGCAGTATTTTGAAGCAAGTGTCCATGGGAAAGATTTCAGTTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ACTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCTGGAGTTTGCTGATGGGACAG
ATTGCAAGCTTCCGGCTCCCTGAGTGCAGAGATATCTGGAGTTTGCTGATGGAAOCA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TGCTCAACATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCTTT
TGATCAACATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCTTT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TCCACAAATGGAACTTGTATTTCAACAACTTGGGATGGCAGAGGAAGGAGATTATCT
TCAACAAATGGAACTTTATCTTCAACAACTTGGGATGGCAGAGGAAGGAGATTATCT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCTCTGCCAGAACACCTTAGGGAAGATGAGTGAAGTCCACCTAACAGTTCTAACAG
GCTATGCCAGAACACCTTAGGGAAGATGAATGAAGTCCACTTAACAGTTATAACAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCATCCCAAGGATAAGGCCAAGCTACAGACCAACATGAGGCTCAGGGCTGGGAAACAG
CTGCTCCCGGATAAGGCCAGTAAACAAACCAACAGAGAAATCAAGCTGGGACACAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTGTCTTGAAGTGGAGGTCAGTGGGGAACCGAAGCCCAATGATTTTGGTTGCTGCTT
CTGTCTTGAAGTGGAGGTCAGTGGGGAACCGAAGCCCAATGATTTTGGTTGCTGCTT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCAACAATGTCAATTCATTCTCCATGACAGGTTCAATTTCTGCAATAGACTTTGT
CCAATGACATGATTTCTTCTCCATGATAGGTACACATTTCTGCAATGGGTCTTTGA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCATCCATAAGTGAACCACTTGAATCTGGGACTATGTGTGCTAGCTCAGATCCTA
CCATCAACAAAGTGAAGTGTGATTTCTGGAGGTACGTATGTGTAGCCGAAATCCCA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GTGGGATGACACTAAGACATACAACTGACATTTGTCTTAACCTCCATTATCAATG
GTGGGATGACACCAAAATGTACAACTGATGTGTCTTAACCTCCATTATCAATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCCTGTATGCAACAGACTGTTATTAAAGCCACAGCCATTCGGCACTCCAAAAAATCT
GTCTGTATACAAACAGACTGTTATTAAAGCCACAGCTGTGAGACATTCAAAAAATCT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTGACTGCAGAGCAGATGGGATCCATCTTCCAGGTCACTGGATTATGCCAGGCAATA
TTGACTGCAGAGCTGAAGGGACCAATCTCTGAAGTCATGTGGATCATGCCAGCAATA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTTTCCTCCAGCTCCATCTTTGGAGCAGAGTCAGGTCCATCCAAATGGAACCTTGG
TTTTCCTCAGGCCCATCTATGGAAGCAGAAATCAGTCCATAAATGGAACCTTGG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGATGAGGAACATCCGCTTTCTGACTCTGCGACTTCACTGTGTGGTTCGGAGGAGG
AAATTAGGAATGTGAGGCTTTCAGATTCAAGCCACTTTATCTGTGTGGCCGAATGAG

03051630 # 030402

rat_cDNA
human 5+3 corrected
mus_cDNA_5
GAGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAGAAGACCAACAT
GTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGGAAGACCGACAT

rat_cDNA
human 5+3 corrected
mus_cDNA_5
TCAGAAACCCATTCAACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAAC
TTAGAAATCCATTTAATGAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTGAATT

rat_cDNA
human 5+3 corrected
mus_cDNA_5
GCTCTGTGGATGGGAACCCCCACCTGAAATTAOCTGGATCTTACCTGAOGGCACACAGT
GCTCTGTGTGATGGTAACCCACCACTGAAATAATCTGGATTTTACCAATGGCACACGAT

rat_cDNA
human 5+3 corrected
mus_cDNA_5
TTGCTAACAGAACACAAATTCOCCTATCTGATGGCAGGCAATGGCTCTCTCATCCTTT
TTTCCATGGACCAAAAGTTATCAGTATCTGATAGCAAGCAATGGTCTTTTATCATT

rat_cDNA
human 5+3 corrected
mus_cDNA_5
ACAAAGCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT
CTAAACAACCTCGGGAGGATGCAGGAAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5
ACATCGAGAACTCATCTGTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGAAC
ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5
CAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTCTCTGATGGGA
CAGGGACAGTAAAGGCCATCAGTGGAGAATCTCTATCACTGCATTGTGTCTCTGATGGAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5
TCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGACAGGCTCAAG
TCCCTAAGCCAAATATCAAAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCTCAAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5
TGGATGGAAATACATACTGCATGAAATGGCAGCTGGTCATCAAGCAACCAACAGCTC
TTAATGGGAATACATATTGCATGACAAATGGCAOCTTAGTCATTAAAGAAGCAACAGCTT

rat_cDNA
human 5+3 corrected
mus_cDNA_5
ACGACCAAGGAAATATATCTGTAGGGCTCAAAACAGTGTGGCCAGGCACTTATTAGCG
ATGACAGAGGAACTATATCTGTAGGCTCAAAATAGTGTGGTCATACACTGATTACG

rat_cDNA
human 5+3 corrected
mus_cDNA_5
TGTCAGTGATGGTGTGTGGCTACCOCTCCCGAATCATAACTAOCCTACCCAGGAACATGC
TTCCAGTAATGATTGTAGCCTACCOCTCCCGAATTACAAATCGTCCACCCAGGAGTATTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5
TCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCCTTGGGAATCCCCAAGCCAA
TCACCAAGGACAGGGGCAAGCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5
AAGTCACCTGGGAGACGCGAAGACACTCOCTGCTCTCAAAAGCAACAGCAAGAAAACCC
AAATCACATGGGAGATGCTGACCACTCOCTTCTCTCAACGGCAAGTAAGAGAGGACAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5
ATAGAAGTGAGATGCTTCACCCCAAGGTACGCTGGTCATTGAGAACTCTCCAAACCTCGG
ATGGAAGTGAGCAGCTTCACTTACAGGTACCOCTAGTCATTGAGAACTCCCAACCTCCG

095433034560

ATTCCGGAGTCTATAAGTGCAGAGCTCAGAACTACTTGGGACTGATTACGGCAACAATT
ATTCTGGGATATACAAATGCACAGCAAGAACCCTCTGGTAGTGATTATGCAGCAACGT

ACATCCAGGTA CTCTGACAGGAAGGGGGAGACTAAAATTCAACAGAGTCCACATOCACA
ATATTCAAGTAATCTGACATGAA-----ATAATAAAGT-CAACAA-----CATCTGGGCA

GGGTTTATTTTTTGGAGAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGA
GATTTTATTTTTTGGAGAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAATTGA

ATACATTACAGTATTAAATTACAATGGACATGCGA--TGA---GACTTGTAATGAAA
ATACATTACAGTATTAAATTACAATGACATGCAAAATAAAAGGACTTGTAATTAAT

GCATTGTGAAGTGA---ACCGAGTCTCTG--TGGHCTCAAAGCAAACCTTTAACTTAA
GCATTATGAAGTGAATGATACTGATTTATTTAATGGATCTCAAACCAAACCTTTAACTTAA

GGCACTTTGATTTTGCCAACAAATAATAACAAACATTAAAGAGAAAAAATGATCCACTAC
GGCACTTTTATTTTGCCAACAAATAACAATAAACAA-----ACATTGAAACGGTTCATAT

GAAATACAAACGGCTAATGCACCTGAATTCT-CAGTAAAAGACCTTCTCTOGCTAAC
AAAATAACAAATGGCTAATGTACCTGAATTTTTCAGTAAAABA--TGAAGTT-CTAAT

AGTTGCCAGCTGCGTGTCTGTTTCTTACCAATGTCACAAACATCGGCACACAGGGTGA
A---CCAGTTGCCTAGTGTCCACCTCTCTATCAATGTTACAGGCATGGCACTCAG---A

ATGGAGTCAACGGGAAGATTAAAGTTTGGGGTCTGTGTAAATCTCAATGTACAAATATTC
ACAGAGACCAATGGAAAATATTAAATCTGC-----AATCT--ATGTATAAATATTT

TGTGCTGGTTTATAAACATTTT-GATAAAACCGAATAAAAAAAAAAAAAAAAAAAAAA
TGT----GGTTTATAAATTTTTTTGTCTAAACCTACAGAAATAAG-----

AAAA
 444-444-444
 444-444-444

(rat_cDNA: SEQ ID NO:7)
(human_5+3 corrected: SEQ ID NO:8)
(mus cDNA_5: SEQ ID NO:9)

Figure 15

rat	MQVGRGVSGLLISLTAVCLVVTFGSRACPRRCACYPTEVHCTFRYLTSIPDGIPANVE
human_5+3_corrected	MKVKGRTITCLLVSEFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTSIPDSIPPVE
mouse_5_corrected	MOKRGREVSCLLISLTAVCLVVTFGSRVCPRRACACYPTEVHCTFRDLTSIPD-GPANVE
rat	RINLGYNLSLRLTENDFDGLSKLELIMLHSHNGIHRVSDKTFSGLSQLQVLKMSYNKVQII
human_5+3_corrected	RINLGYNLSLRLMETDFSGLTKELEIMLHSHNGIHTIPDKTFSDLOALQVLKMSYNKVRKL
mouse_5_corrected	RVNLGYNLSLRLTENDFSGLSKLELIMLHSHNGIHRVSDKTFSGLSQLQVLKMSYNKVQII
rat	RKDTFYGLGSLVRLHLDENNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTEVSLSYLQ
human_5+3_corrected	QKDTFYGLRLSLRLHMDENNIEFINPEAFYGLNRLVHLEGNRLTKLHPDTEVSLSYLQ
mouse_5_corrected	EKDTLYGLRLSLRLHLDENNIEFINPEAFYGLTLRLVHLEGNRLTKLHPDTEVSLSYLQ
rat	IFKTSFIKYLELSDNFLTSLPKEMVSYMPNLESILYHGNPWTCDCHLKWLEWMOGNPDI
human_5+3_corrected	IFKISFIKLYLSINFLTSLPKEMVSYMPDLDSLYLHGNPWTCDCHLKWLESDWIO--PDV
mouse_5_corrected	IFKTSFIKLYLYDNF-TSLPKEMVSSMPNLESILYHGNPWTCDCHLKWLEWMOGNP--
rat	IKCKKDRSSSSPQQCPLCMNPRIKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTOEDNG
human_5+3_corrected	IKCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSS
mouse_5_corrected	-----
rat	SASTSPQDFIEPFGSLSLNMTXXSGNKADNVCSIQKPSRTSPATTEENDYTMINASFTST
human_5+3_corrected	SAFISPGGFMAPFGSLTINMTDQSGNEANMVCSIQKPSRTSPATTEENDYIVLNTSFTST
mouse_5_corrected	-----
rat	NIVCSVDYNEHQPVWQLALYSDSPLILERKPKLTETPSLSSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIDYGHQIPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEA
mouse_5_corrected	-----
rat	DVRADFFWFQOEKIVLQLNRTATTLSLQIQFSTDAQIALPRAEMRAERLKWTHILMMNN
human_5+3_corrected	DLRADPSWLMQDQISLQLNRTATTFTSLQIQYSSDAQITLPRAMRPVKHKWTMISRDNN
mouse_5_corrected	-----
rat	PKLERTVLVGGTIALSCFGKGDPSPHLENLLADGSKVRAPYVSEDGRILIDKNGKLELQM
human_5+3_corrected	TKLEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
mouse_5_corrected	-----
rat	ADSPDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDELPCISTGV
human_5+3_corrected	ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYOENGIHHTVFIGETLDELPCSTGI
mouse_5_corrected	-----
rat	PIASISWILPGNTVFSQPSRDAQILNNGTLRLQLQVTPKDGHYQCVANPSCADFFSSFY
human_5+3_corrected	PIASISWVIPGNVLYQSSRDKKVLNNGTLRLQLQVTPKDGYYRCVAPNPGVDFLIFVY
mouse_5_corrected	-----
rat	SVQKKGQRNVHEDREAGGSGLGEPMSSVSLKQPSILKLSASALTGSEAGQVSGVHRKNK
human_5+3_corrected	SVKMKGORPLEHDGETEGSGLDESNIPIAHKEPPGAQLRTSALMEAEVGHKTSSTSKRNN
mouse_5_corrected	-----
rat	HRDLIHRRRGDSILRRFREHRRQLPLSARRIDPQWAALEKAKKNSVPPKQENTTVKVP
human_5+3_corrected	YRELTLQRRGDSILRRFREHRRHFPSPARRIDPQWAALEKAKKNAMPDKRENTTVSFP
mouse_5_corrected	-----

20190923 09:56:56

rat PLAVPLVELTDEEKDASGMIPPEEFMVLTKASGVPGRSPTADSGPVNHCMTSASGT
human_5+3_corrected PVVTQLNPNGEEDDSSGMLALHEEFMVPATKAINLPARTVTADSRITSDSMNTNINYGT
mouse_5_corrected -----

rat EVS-TVNPQTLOSEHLPDFKLFVTVNGTAVTKSMNPSLASKIEDTTNQNPIILFP---SV
human_5+3_corrected EFSPVYNSQILPPEEPTDFKLSTAIKTAMSKNINPTMSSQIQGITNQHSSSTVFILLGA
mouse_5_corrected -----

rat AEIRDSAQAGRAS--SQSAHPVTGGRMATYGHNTYSSFTSKASTVLQPINPTESYGPQI
human_5+3_corrected TEFQDSQMGGRGRHFQSRPPITVRIMIKOVNVKMLSSTTNKL---LLESVNTTNSH--QT
mouse_5_corrected -----

rat PITGVSRPSSSDISSHTADP8FSSHPSGSHTTASLLEHIPPNNNTGNFPLSRHLGRERT
human_5+3_corrected SVREVSEPRENHFYSHTTQILSTSTEPSDPTAHSQFPIPRNS-TVNIPLFRFRGRQRK
mouse_5_corrected -----

rat IWSRGVKNPHRTFVLRHRHRTVREPAIKGPANKNVSVQVATEYPMCHTCPSAEGLTVA
human_5+3_corrected IGGRGRIISPYRTPVLRHRHYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRERLTTA
mouse_5_corrected -----

rat TAALSVPSSSHSALPKTNVGVIAEESTTVVKKPLLLFKDKONVDIEIITTTTKYSGGES
human_5+3_corrected TAALSFPSSAAPITFPKADLARVPSEESTTLVQNPILLLENKP--SVEKTPPTIKYFRTET
mouse_5_corrected -----

rat NHVITEASMTSAPTSVSLGKSPVDNSGHLMPGTIQTGKDSVETTPLESPILSTP--SIP
human_5+3_corrected SQVTPTGAVMTYAPTSIMEKTHKVNASYPRVSSSTNEAKRDSVITSSLSGAIKTPMTII
mouse_5_corrected -----

rat TSTKFSKRKTPLEHQIFVNMOKKEGMLKNFYQFGLQKNPAKLPKXIAPLLPTGQSSPSDST
human_5+3_corrected AITRFSRRKIPWQONFVNMHNPKGFLRNQHKVSLQKSTAVMLPKTSPALPQRSSPFHFT
mouse_5_corrected -----

rat TLLTSPPPALSTTMAATONKGTVEVVGARSLSAGKKQ-PFTNSSPVLPTIEKRNTLINE
human_5+3_corrected TLSTSVMOIPSNLTATAHTTTKTHNPG-SLPT-KKELPPFPLNFMPLPSIISKDSSTKSI
mouse_5_corrected -----

rat ISTETPT-VTSPTATASVIMSETQRTSRKEAKDQIKG-P-RKNRNNTTTPROVSGYSAY
human_5+3_corrected ISTQTALPATTPTFPASVITYETQTERSRAQTIQREQEPQKKNRTDPNISPDQSSGFTTF
mouse_5_corrected -----

rat SALTTADTFLAFSHSPRODDGCVSAVAYHSTTS--LLAITELFEKTYTTLGNTTALETT
human_5+3_corrected TAMTP-PALAPTHSPENTTGISSTISFHSRTINLTDVIEELAQASTQTLKSTIASETT
mouse_5_corrected -----

rat LLSKSQESTTVKRAS-DTP-PFLLSSGAPFVPTSPFPPTKGVVTDGKVTSFAQMTSNRV
human_5+3_corrected LSSKSHQSTTTTKASLDTPIPPFLSSSATIMPVPISPFPTQRAVTDTRGDSEFRIMTNTV
mouse_5_corrected -----

rat VTIYESSRENTDLQQP8AEASPNPEIITGTTDSPNLFPSTSVPALRVDPKNSKWKESP
human_5+3_corrected VKLHESSRN--LQMPSSQLEP-----LTSSTSNLLHSTPMPALTTVKSNKLTSPSP
mouse_5_corrected -----

rat WPEHKYOLKSYSETIEFKRPVSMSPHLSPLEASTHASEWNTQKHAESVFDKPKGQNP
human_5+3_corrected WAEYQFWHKPYSDIAEKKGKPEVSMLATTLGLSEATTLVSLWDGQKNTKSDFKPKPVQEA
mouse_5_corrected -----

00041630 030402

[illegible][illegible]

rat	PITGVSRPSSSDISHTTADPSFSSHPGSGHTTASSLFIHPANNNTGNFPLSRHLGRERT
human_5+3_corrected	SVREVSEPRHNHFYSHTTQILSTSTFPSPDEHTAAHSQFPIPNNS-TVNIPLFRFRGRQRK .: **.* ..: **** * *;..**.* * * ****. * *.* *;:***.*
rat	INSGRVRKPHRTFVLRHRHRTVRPAIKGPANKVNSQVPATEYPMCHTCPSAEGLTVA
human_5+3_corrected	IGGRGRIISPYRTFVLRHRYSIFRSTTRGSSEKSTAFSATVLNVTCLSCLPRLTTA * .***. .*:*****; .*: .*:..*..: ..** * : * . * **.*
rat	TAALSVPSSSHSLPKTMNVGVIAEESTTVVKPPLLKPKQNVDEIITTTTKYSGGES
human_5+3_corrected	TAALSFPSAPITFPKADIARVPSEESTTLVQNPILLLENKP--SVEKTPPTIKYFRTET *****,*:: :***: . * ;*****;*;*****;:~* .:* *.* ** *
rat	NHVIPTASMTSAPTSVSLGKSPVDNSGHLMPGTIQTGKDSVETTPPLSPLSTP--SIP
human_5+3_corrected	SGVTPTGAVHTYAPTSLNMEKTHKVNASYPRVSTNEAKRDSVITSSLSKATKPKMTII .:* ** * ** ****;.: * : *!..: ..* : : *** *;.*.....* : *
rat	TSTKFSKRKTPHLQIFVNNQKKEGMLKNPYQFGLQKNPAAKLPKLAFLPTGQSSPDSST
human_5+3_corrected	AITFSRRIKIPWQCNFVNNHNPGRRLRNOHKVSLQKSTAVMLPPTSPALFORQSSPHFT ; *::** * ; * ****;: * *.* :!..***..* . *** : * ** **** . *
rat	TLLTSPPPALSTTMAATONKGTENVVSGARLSAGKQPTNSSPVLPTISKRSNTLNFL
human_5+3_corrected	TLSTSVMQIPSNLTITTAHTTTKTHNPG-SLPTKKELPFPPLNPMPLPSIISKDSSTKSI ** ** *.:***:!!: *.. . **.: * : ** . *:*** ** *.* ::
rat	STETPT-VTSETATASVIMSETQRTSKEAKDQIKG-P-RKNRNANTTTPROVSGYSAYS
human_5+3_corrected	STOTALPATPTTFPASVITYETQTERSRAQTIQREOEPQKQNRDTPNISFDQSSGFTPT **.*. .:*** .**** ** *; . * : * ;***.: * : * * *::: :
rat	ALTADTFLAFSESPRODGGNVSAVAYHSTTS--ILAITELFEKYTOTLIGNTTALETTL
human_5+3_corrected	AMTP--PALAFTHSPPEMTTGISSTISFHSRTLALTVIEELAQASTQTLKSTIASETTL *.*. .:***:*** : * *;::** * . * ** : **** . * * ****
rat	LSKSOESTTVKRAS-DTP-PPLSSGAPPVPTSPPPPTKGVVTDKSVTSAFQMTSNRVV
human_5+3_corrected	SSKSHQSTTRKASLDTPFPPELSSATLMPVPISPPTQRAVTDTRGDSHFRLMTNTVV ***::***;:* ** *;*;***.* . *.* .****: .***: * *; : * **
rat	TIYESSRHNTDLOQPSAEASNPETITGTDSPLNLPSTSVPALRVDEKQNSKNKPSFW
human_5+3_corrected	KLHESRHN--LQMPSSQLEP-----LTSSTSNLHSTMPALTTVKSQNSKLPSPFW .:***** ** *::.* *.*.***: **.* ** . * **** .****
rat	PEHKYQLKSYSETIEKGRPAVSMSPHLSLPEASTHSHWNTQKHAESVFDKPKQNP-
human_5+3_corrected	AEYQFWHKPYSDIAEKKGKPEVSMLATTLGLSEATLVSDWDGQNTKESDFDKPVQEAT .*::: *.*: ****:* ** , .*.**.*. *.*: **::** ***** *;.
rat	TSKHLPEVSLPKTILKKPRIIGGKAASFTVPANSDFLPCAVGDPLPIHNTVSSGXE
human_5+3_corrected	TSKLLPFDLSLSRYIFEKRIYGGKAASFTIPANSDAFLPCAVGNPLETIHNTVVS-GLD *** **; *.*; :.:***;*****;*****.*****;*** ***** * :
rat	ISQGTQKSRFEVLNPGTLSIQRVSIQDRGQYLCSAFNLGVDHFEVLSLVVFPARILDR
human_5+3_corrected	LSRGNQNSRVQVLPNGTLSIQRVSIQDRGQYLCASNLFGTDHLVTLVSVVSPRIILER :*.*.*.*;*****;*****.***** * :.*.*;*** ** .***.*
rat	HVKEITVHEGSTVELKCRVEGMPRTVSNILANQTVVSETAKGRKVVVTPDGTILIIYNL
human_5+3_corrected	RTKEITVHSGSTVELKCRAGRPSPPTVTWILANQTVVSESSQGSROAVTVVDGTILVLENL :.****** *****.* * **;*****;***:***. ** *****:.*
rat	SLYDRGFYKCVASNPQGDSLLVKIQVITAPPVILEQKQAIIVGLGSSLKLPCTAKGTP
human_5+3_corrected	SIYDRGFYKCVASNPQGDSLLVKIQVIAAPPVILEQKQVIVGTWGESLKLCTAKGTP *;*;*****;*****.*****.*****;***.*. * *****
rat	QPSVHWVLYDGTCLKPLQLTHSRFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVV
human_5+3_corrected	QPSVHWVLYDGTCLKPLQLFTNSKLELFSNGTLYIRNLASSDRGTYECIATSSSGSERRVV ***;*** *****;***:***: .***.*****.*.* *****
rat	ILTVEGETIPRIETASQKNTEVNLGKILLNCSATGDPKPRIIWRLPKSAVIDQNHKMG
human_5+3_corrected	MLIMEERVTSPIEAASQKTEVNEGDKILLNCSATGEPKQIMWRLPKSAVIDQ----G ;*.** * ****;*** **;.*;*****;***:***:*****.* *

[illegible]

(rat: SEQ ID NO:13)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYPTEVHCTFRDLTSIPDGPANVER
VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQIIE
KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDFTVSLSYLQIF
KTSFIKXLYLYDNFTSLPKEMVSSMPNLESYLHGPNWTCDCHLKWLSEWMQGNP
(SEQ ID NO: 15)

Figure 18

MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTIPDSIPPNVE
RINLGYNLVRMLMETDFSGLTKELELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDFTVSLSYL
QIFKISFIKFLYLSDNFTSLPQEMSYPDLDSLLYHGPNWTCDCHLRWLSDWIQPDVI
KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
AFISPQGFMAPFGSLTLNMTDQSGNEANMVCISQKPSRTSPIAFTEENDYIVLNTSFSTFL
VCNIDYGHIQPVWQILALYSDSPILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
ADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEVRPVKHKWTMISRDNNTK
LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
ADSFDTGVIYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGIP
DASISWVIPGNNVLYQSSRDKKVLNNGTLRILQVTPKDQGYRCVAANPSGVDFLIFQV
SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVKGHTSSTSKRH
NYRELTQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALEKAKKNAMPDKRENTTV
SPPPVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRTISDSPMTNIN
YGTEFSPVVNSQILPPEEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSSTVFPLLLG
ATEFQDSDQMGRGREHFQSRPPTVTRTMIKDVNVKMLSSSTTNKLLLESVNSHQTSVREV
SEPRHNHFYSHTTQILSTSTFPSPHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGGRGRIIS
PYRTPVLRHRHYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI
TFPKADIARVPSEESTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
SIPMEKTHKVNASYPRVSSTNEAKRDSVITSSLSGAITKPPMTIIAITRFSRRKIPWQQNFV
NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNLTLT
AHHTTTKTHNPGSLPTKKELPFPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
TQTERSRAQTIQREQEPQKKNRTDPNISPQSSGFTTPTAMTPPALAFTHSPPENTTGISST
ISFHSRTLNLTDVIEELAQASTQTLKSTIASSETTLSSKSHQSTTTTRKASLDTPIPPFLSSAT
LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLMQMPSSQLEPLTSSTSNLL
HSTPMPALTTVKSQNSKLTPSPWAEQFWHKPYSDIAEKGGKPEVSMLATTGLSEATTLV
SDWDGQKNTKKSDFDKKPVQEATTSKLLPFDLSRYIFEKPRIVGGKAASFTIPANSDAF
LPCEAVGNPLPTIHWTRVSGLDLSRGNQNSRVQVLPNGTLSIQRVEIQDRFQYLCSASNL
FGTDHLHVTLVSVSYPPRILERRTKEITVHSGSTVELKCRAGEGRPSPTVTWILANQTVVSE
SSQGSRQAVVTVDGTLVLHNLSIYDRGFYKCVASNPGGQDSSLVKIQVIAAPPVILEQRR
QVIVGTWGESLKLPTAKGTPQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSSTGSERRVVMLTMEERVTSPIEAASQKRTEVNFGDKLLLNC SATGEP
 KPQIMRLPSKAVVDQGSWIHPNGSLFIGVTEKDSGVYLCVARNKMGDDLILMHVSLR
 LKPAKIDHKQYFRKQVLHGKDFQVDCASGSPVPEISWSLPDGTMINNAMQADDSGHR
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIQSNKTN
 KRIKAGDTAAVLDC EVTGD PKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
 EYVCVARNPSGDDTKMYKLDVVS KPPLINGLYTNRTVIKATAVRHSKKHFD CRAEGTP
 SPEVMWIMP DNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGE SVLVVQLE
 VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSDGNPPPEIIWILPNGTRFSNGPQSY
 QYLIASNGSFIISKTTREDAGKYRCAARNKVGYIEKL VILEIGQKPVILTYAPGTVKGISGE
 SLSLHC VSDGIPKPNIKWTMPSGYVVD RPQINGKYILHDNGTLVIKEATA YDRGN YICKA
 QNSVGHTLITVPVMIVAYPPRITNRPPRSIVTRTGA AFQLHCVALGV PKPEITWEMPDHS
 LLSTASKERTHGSEQLHLQGT LVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
 NO: 16)

Figure 19

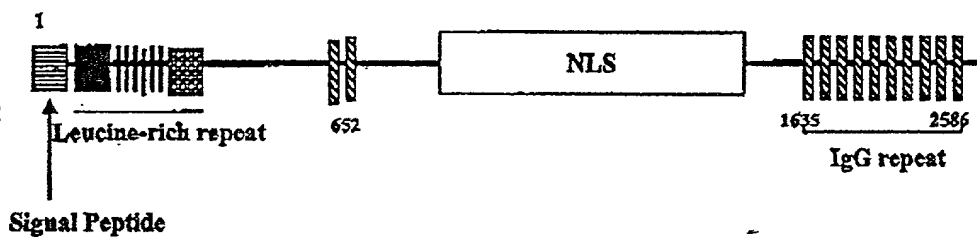


Figure 20

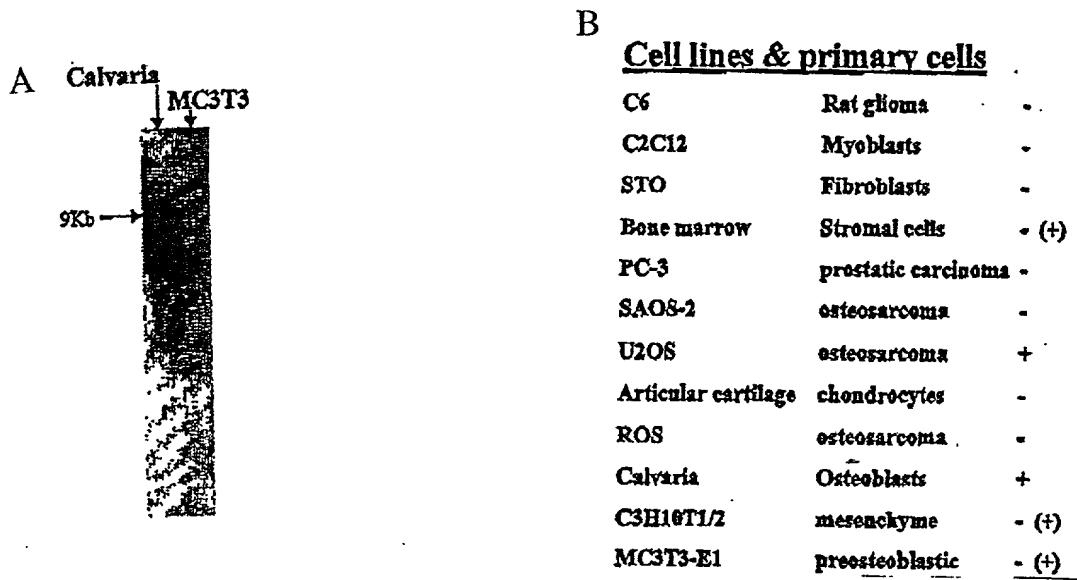


Figure 21

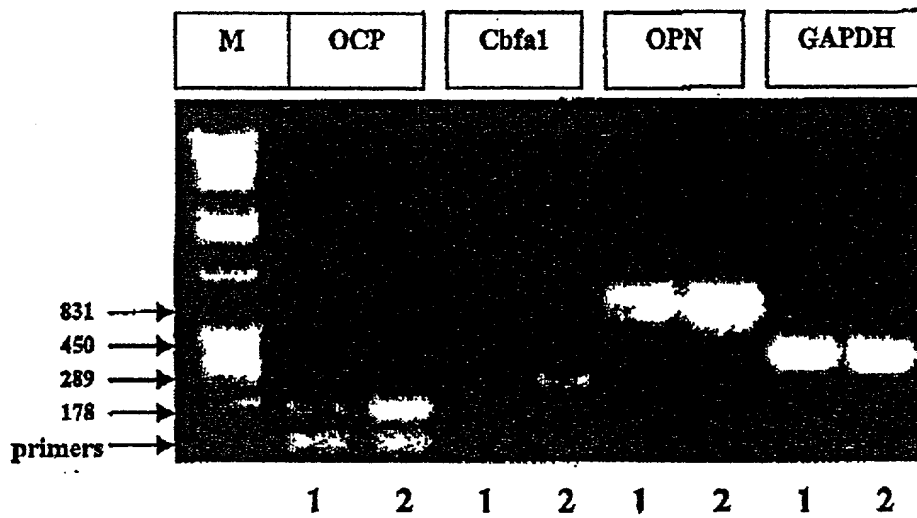


Figure 22

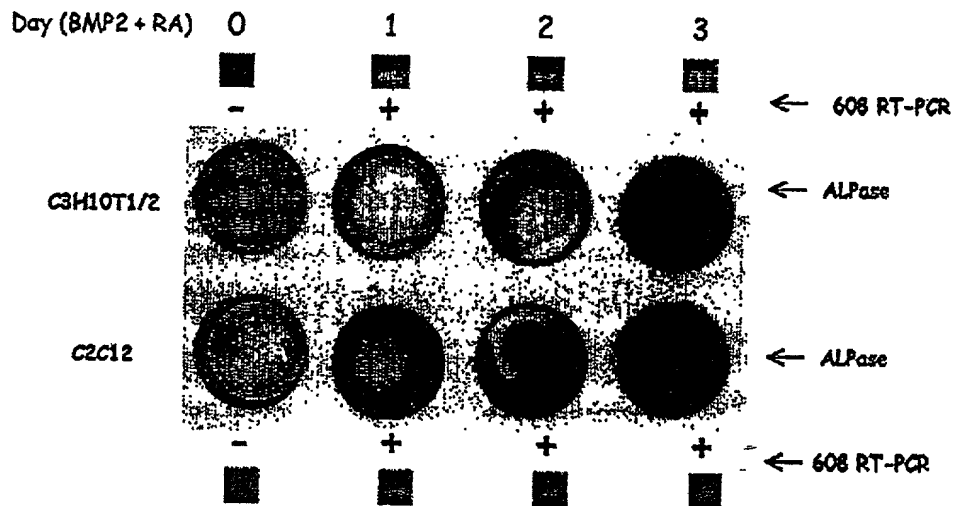


Figure 23

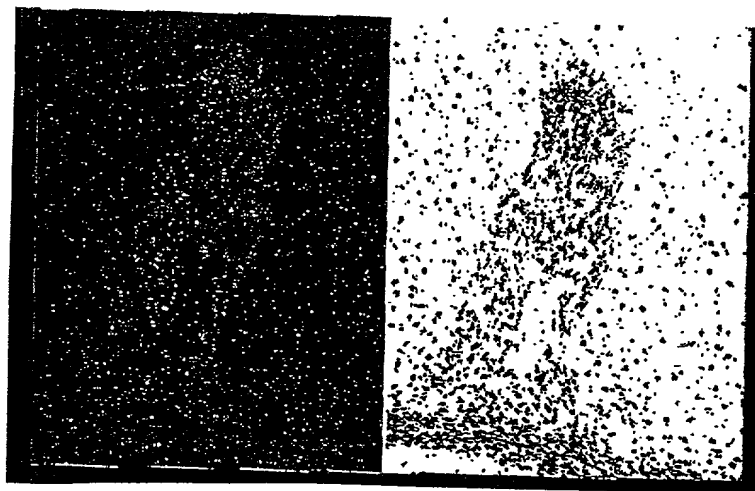


Figure 24

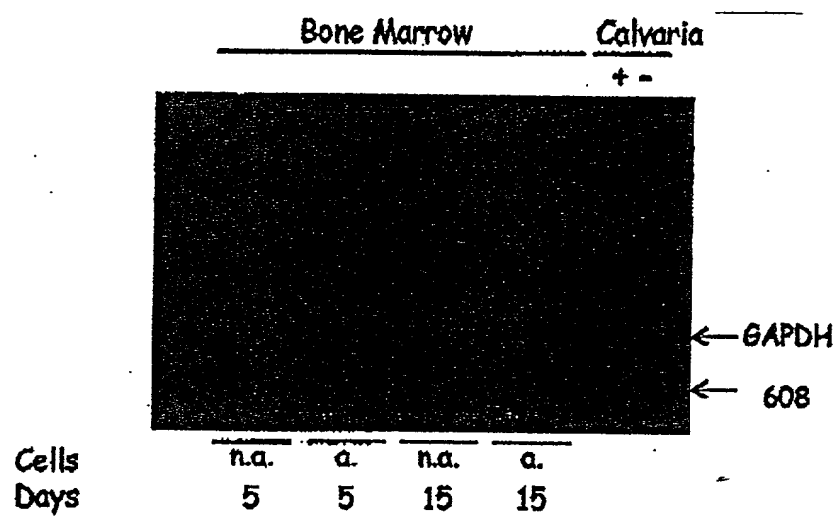


Figure 25

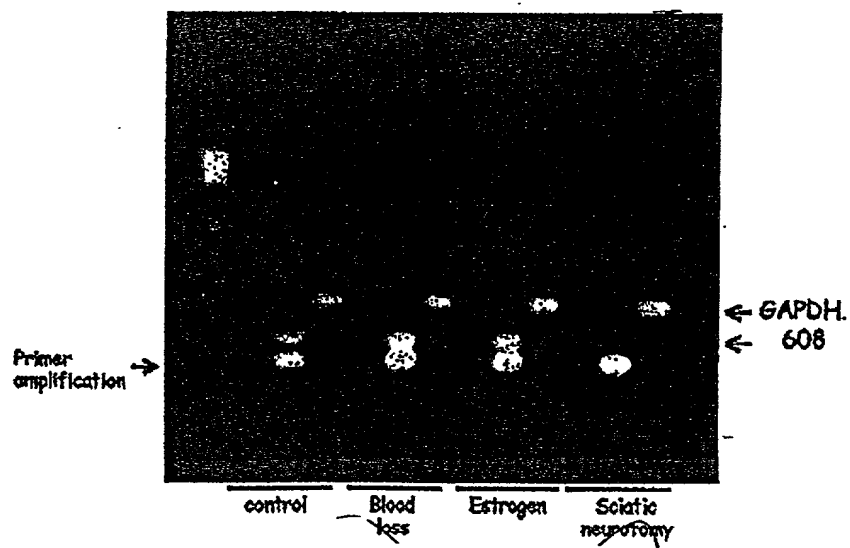


Figure 26



204030 007550

Figure 27

A

B

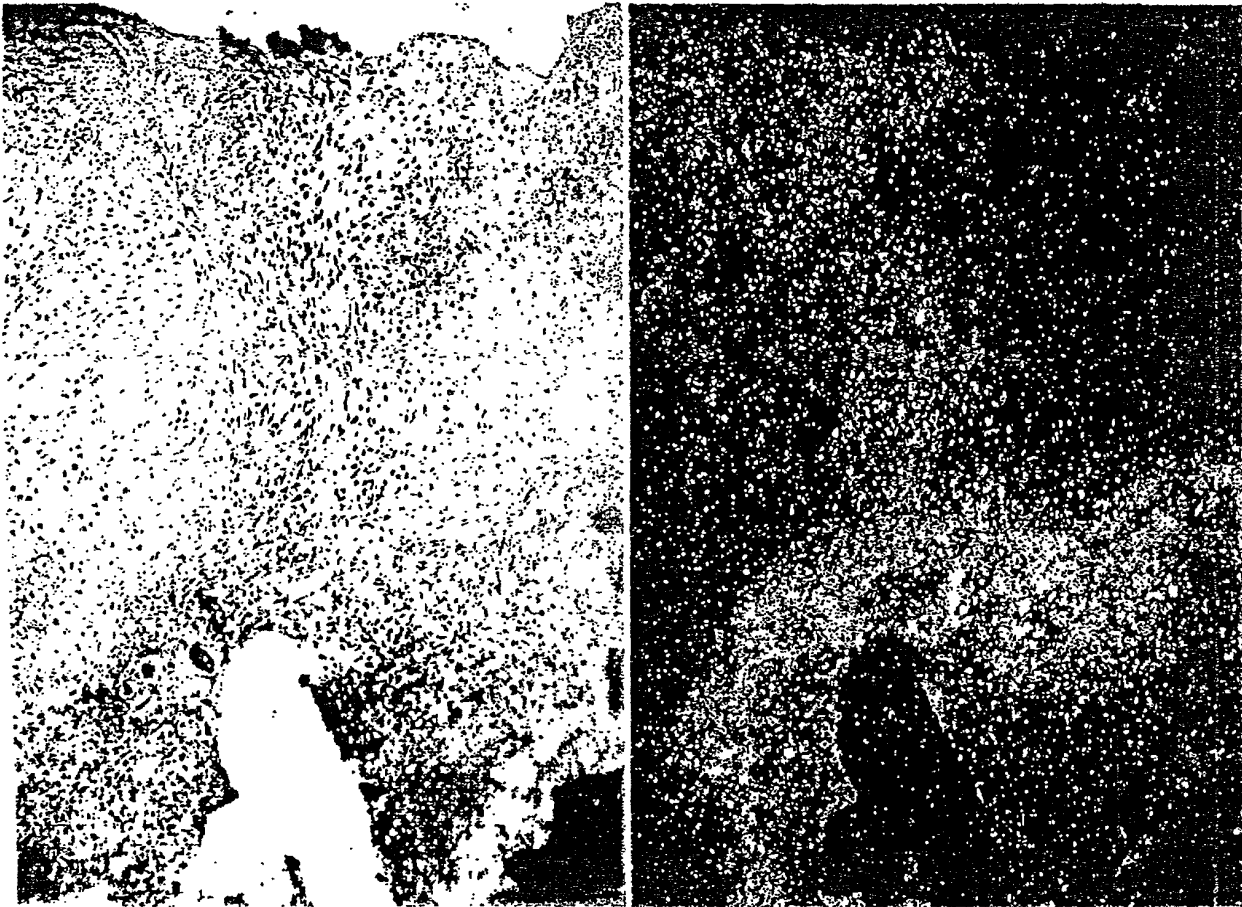


Figure 28



20100101 000000

Figure 29

A

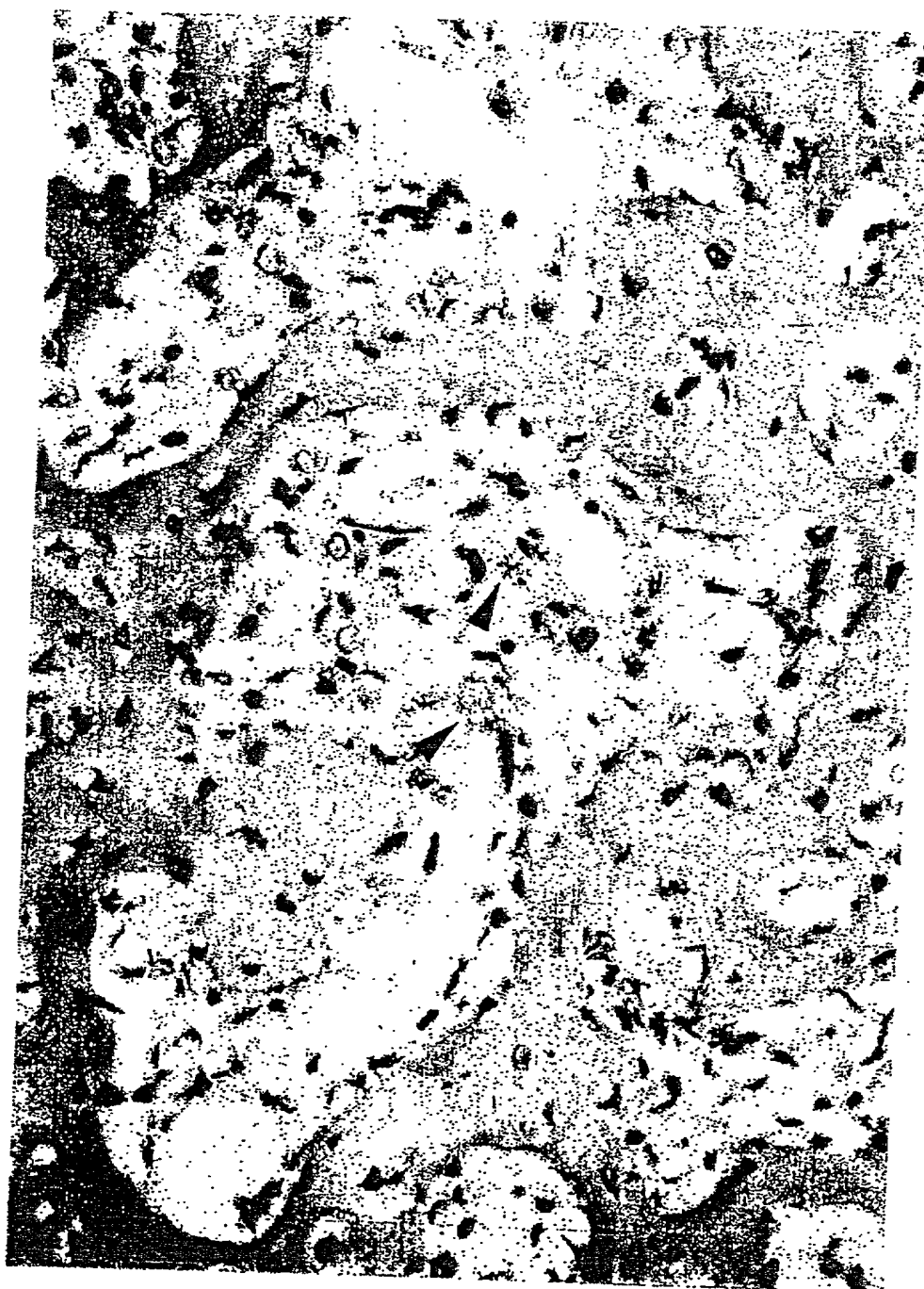


B



00991630 020402

Figure 30



204060 037660

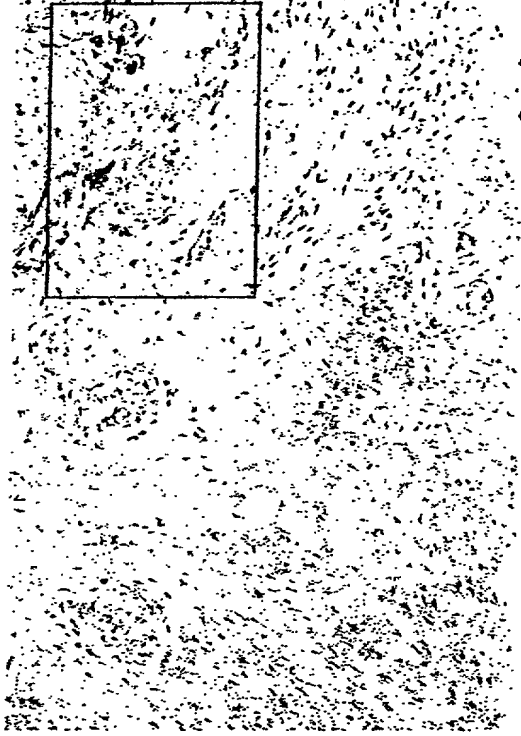
Figure 31



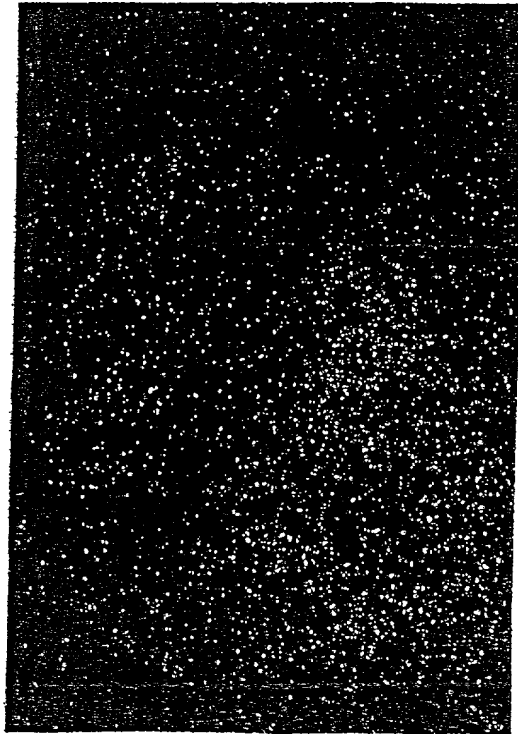
2040ED-0E9T660

Figure 32

A

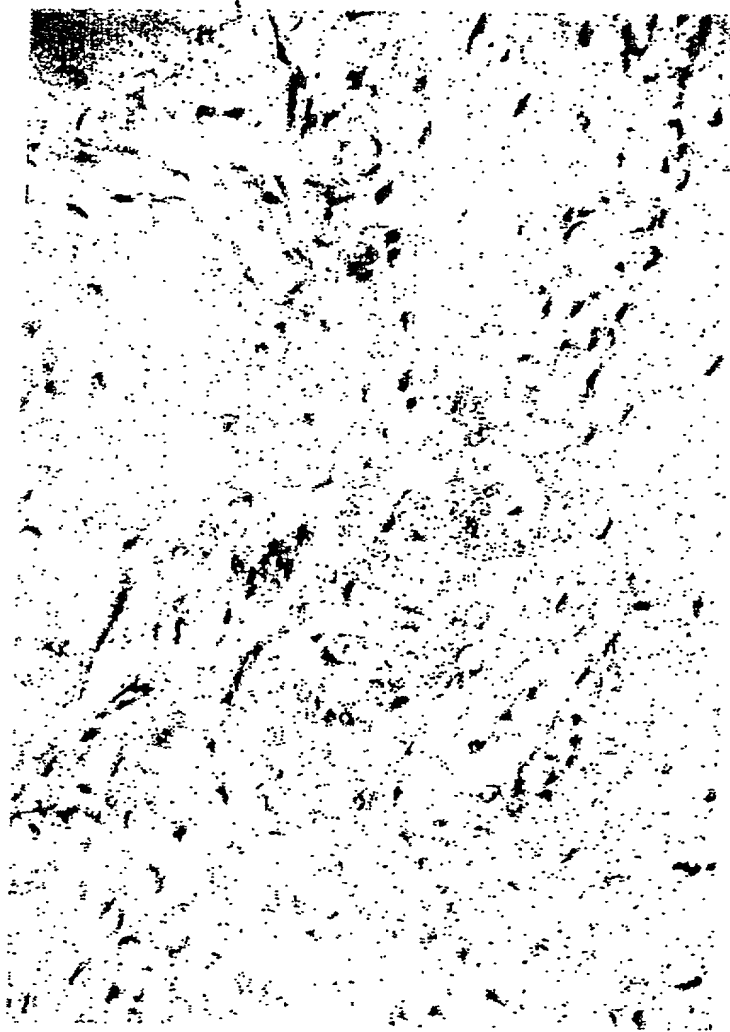


B



207060 0501550

Figure 33



204050 0507550

Figure 34

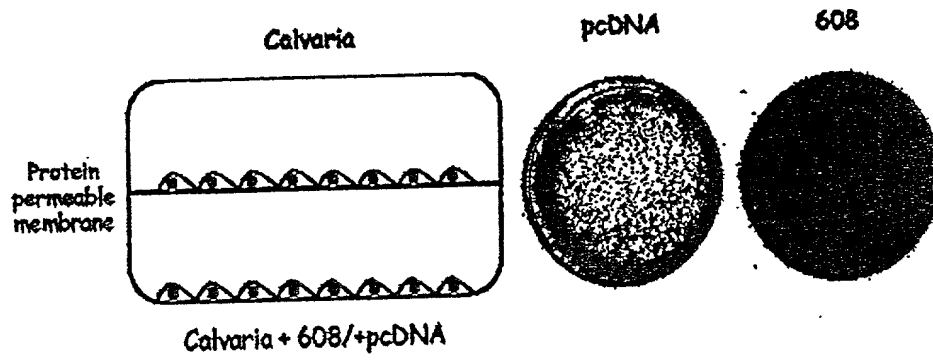


Figure 35

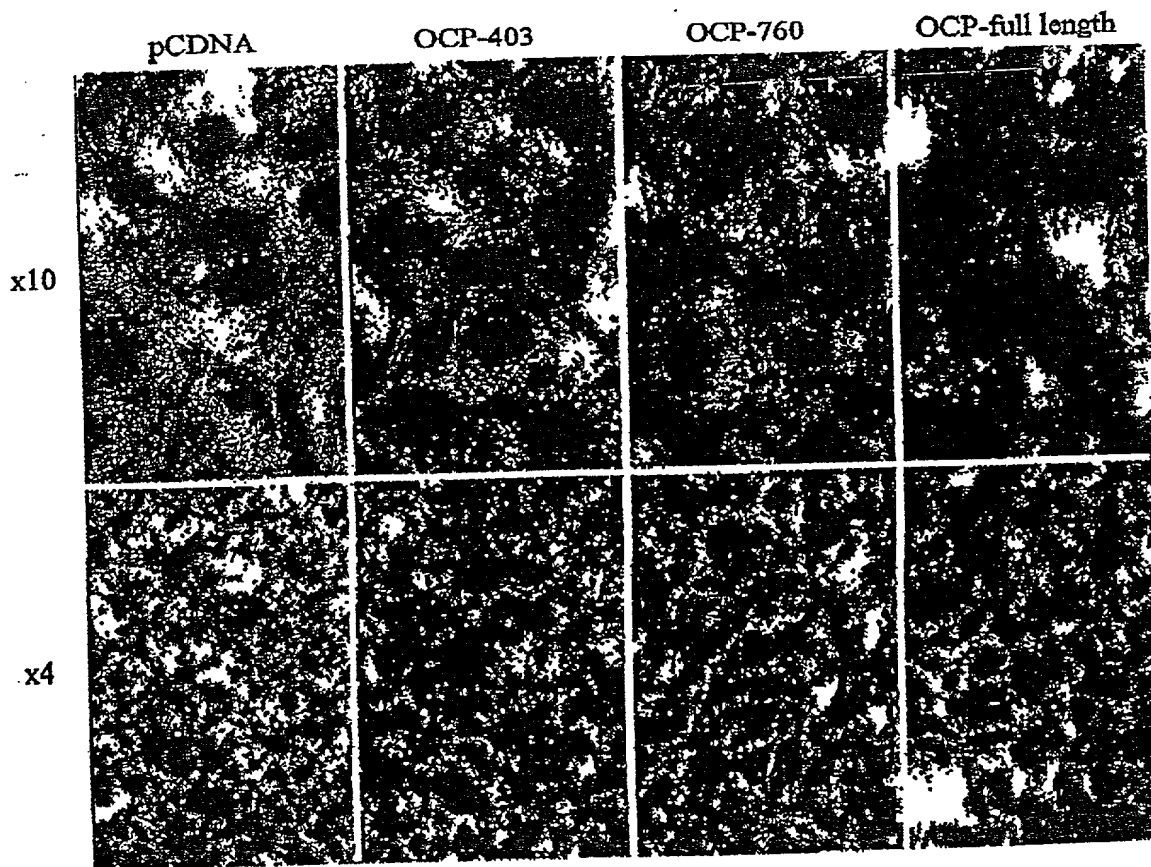


Figure 36

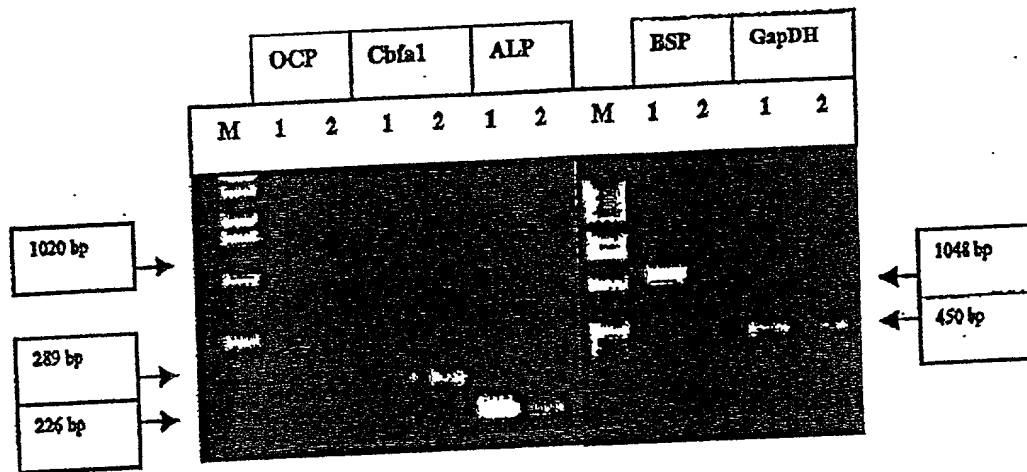


Figure 37

x 4

x 10

pCDNA
ROS stable line

OCP
ROS stable line



Figure 38

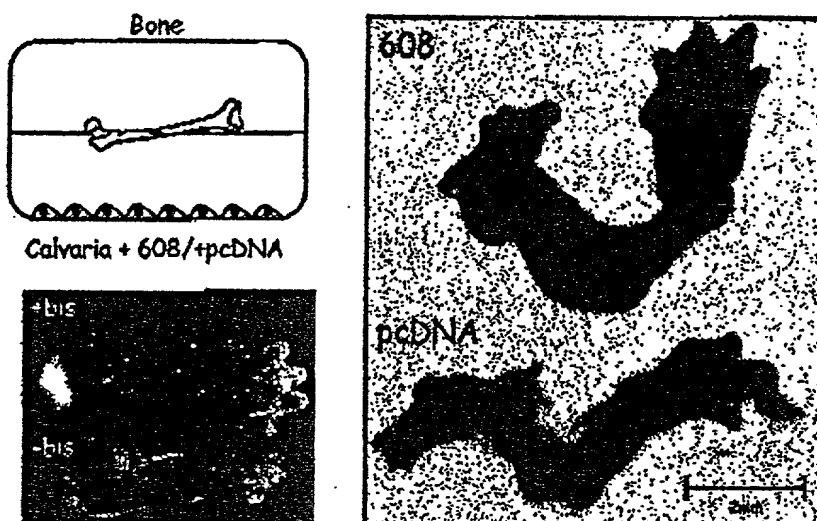


Figure 39

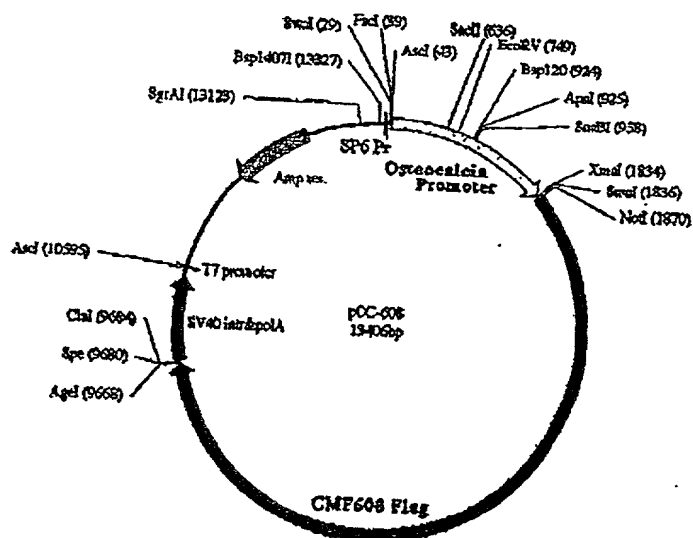
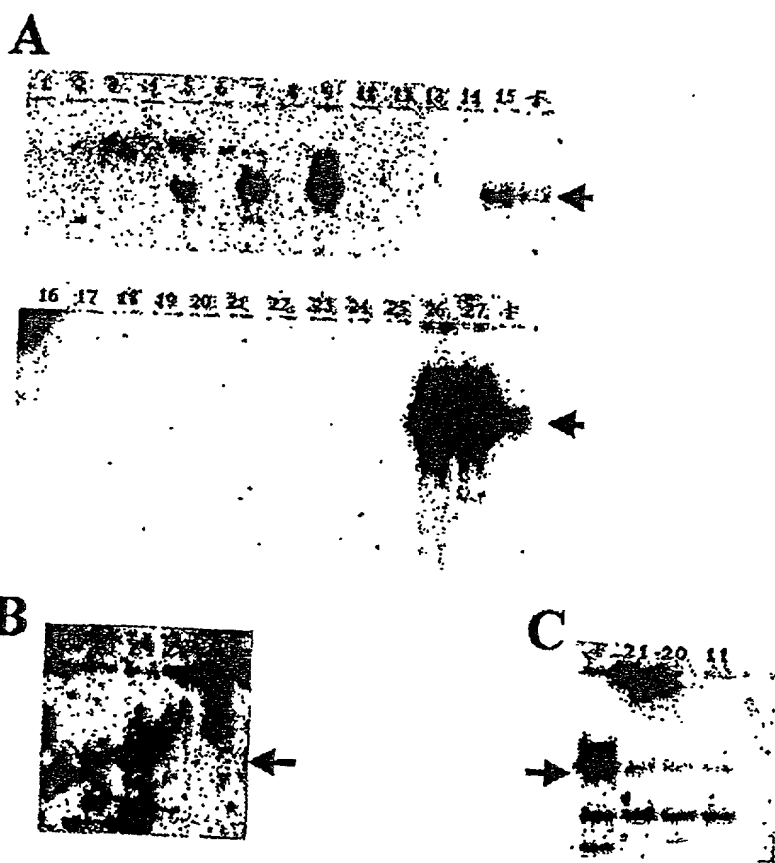


Figure 40



09091630 030492
2074020 030492

Figure 41

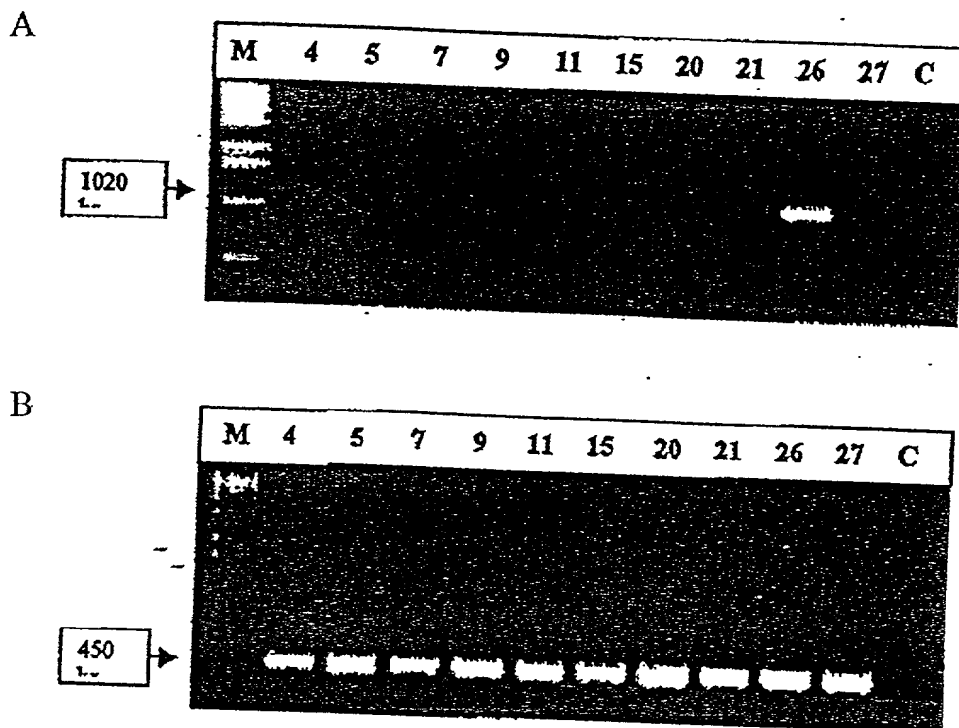
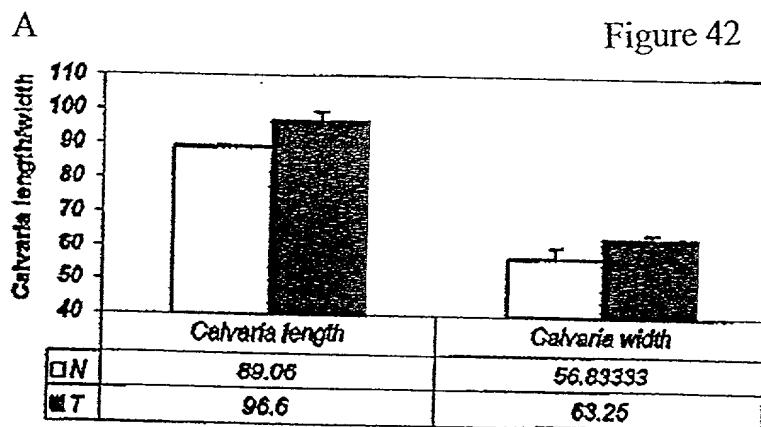


Figure 42



B

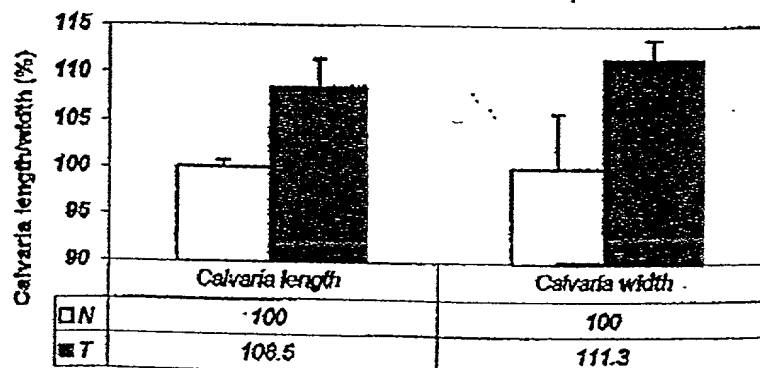


Figure 43

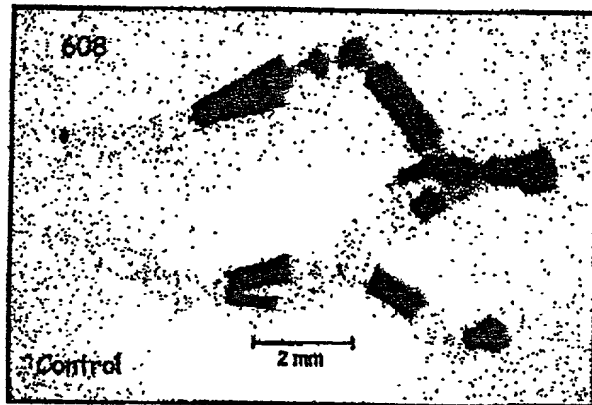


Figure 44

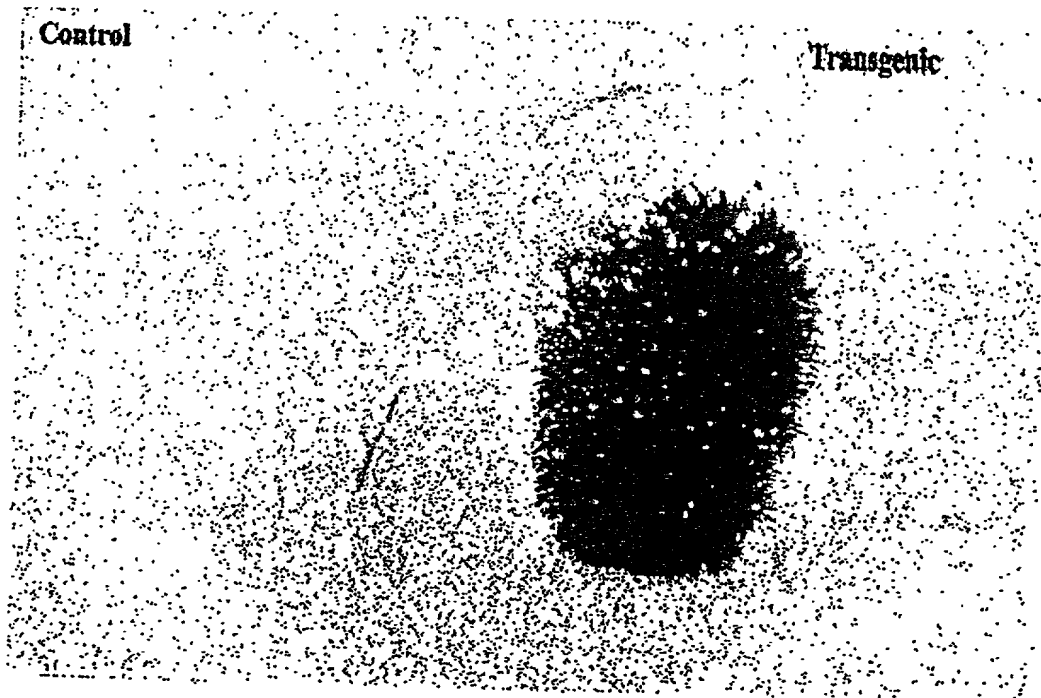
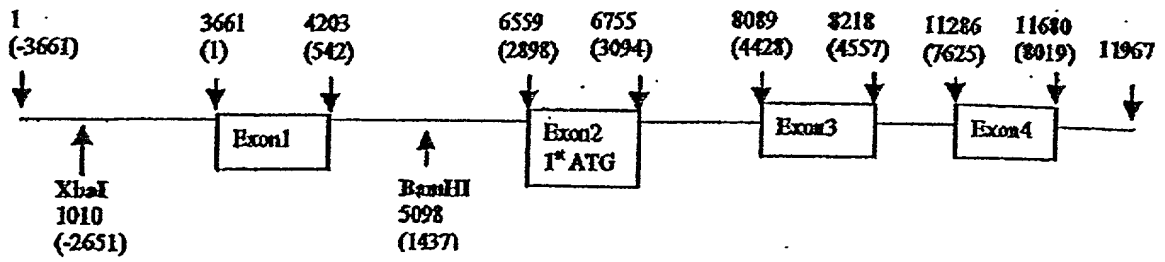


Figure 45



Clone 14C10 /15E11

T7 orientation in pKS

Figure 46

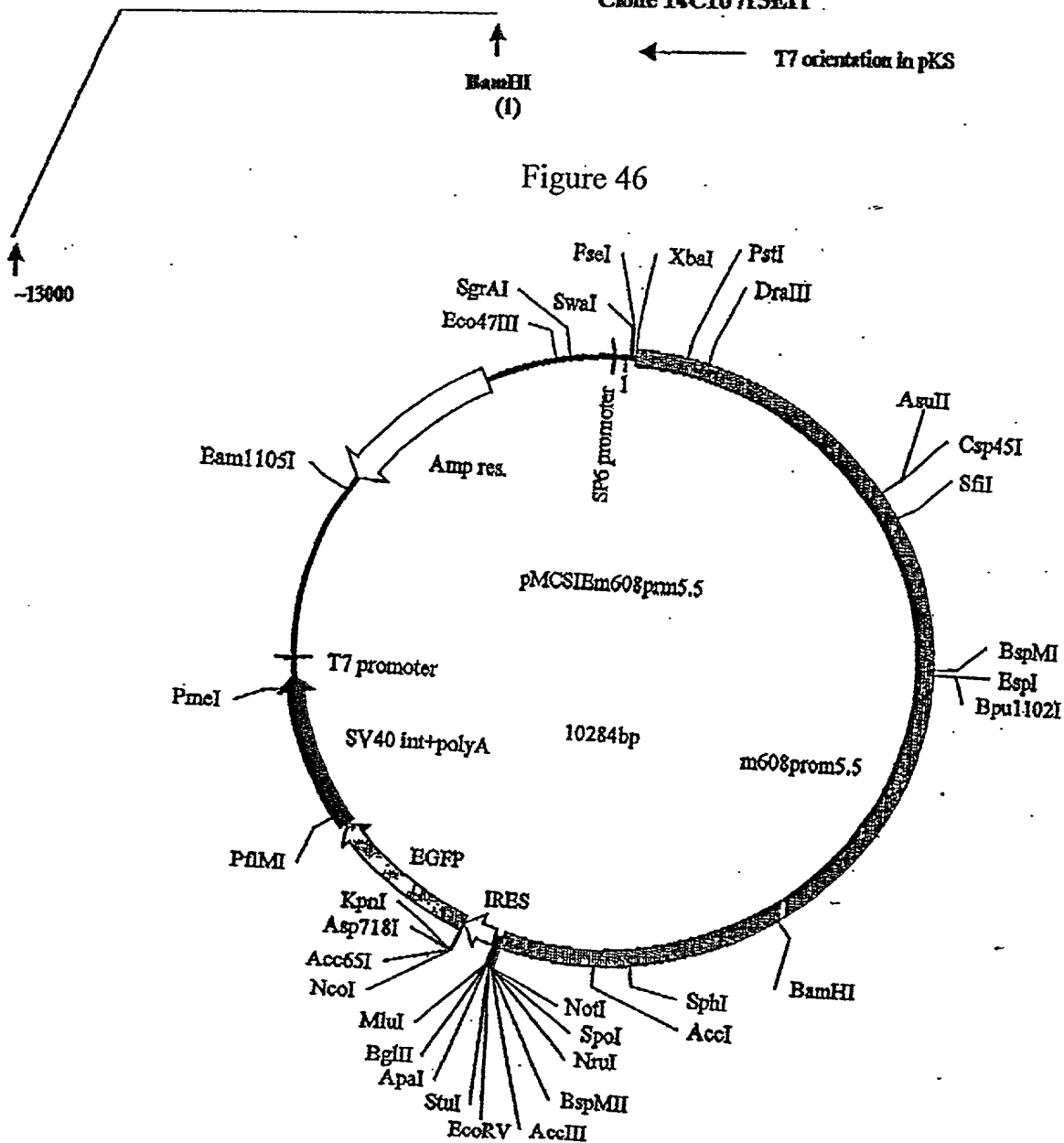


Figure 47

↓ (XbaI)

TCTAGAAGTAAATGATCCTGAGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGC
AATCATCTCTCTGTGGTTTGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATA
ACCTACCGAATACAGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCT
GTAGCTGCCCAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCCGACAT
CTTAAGAATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCGTGCCACACAGGTCTOCAG
GGTTAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTACCTTGATCATTAAATGACTAGAA
ATCACTTGGTGATTAGCAACTGGATATGGAATATTACTAATTTGTACCCAAGCCAGGCCAC
CTCAGCTTTGGCAGCTCCATTTCATTCTGTGGAGCCAGTCACGTGGGTTTGAATCAACTGT
ACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTTCACAAGTCTTTTTA
GAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATG
TGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAATAAACTGACTTTAAAAAATAA
AGCTTATGATACTTGTACAGAGTAAATCTTCCATAAATATCATCTGCATTTATAAATTAT
TTTCATAATCCATCAATTAACAACTTTAGAAATTTTGTTAACACAAAGATCCCTAGGCC
CTGCCCTAGGATGGTCTGTATGGTGGGCTGAGAGATGGAGCTTAAGAACTTACTTGCTCC
AGGAGCACATCTTCAAGACATCTGCCTCAAAAACATTTATCCCAAATGCTCATCAAAGGCTC
ACTCACATGTGCTTCAACCACAGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGT
GGAAGCCTGCTAGAGGAACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTC
ATAACCTCAACCACAACAAACAAACGACAACAAACCCATAAAAAATTATCAGCGCAAAAT
GAACCTAAGCCATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTCTTAAT
AGTCAAAATATGCAGAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATA
CTTGTAGATTAGAGGACACAAAGCAACAACTACAGAGTGAAGTAATCCAGAGACTTAATG
TATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGTATCTT
ACTAACACACACACTTACATAGAGAGATAAAGTGATAGATACGTTTGTITTTATCTTCATG
TAGCTGATAATTTTCATATTGTACACCTCAACATAGATAACCAACAAAGAGGAAGAGGAT
AGGTGCCTCTCCAGGGCGGAAGAGTACATTGCAAAAGTCAGACACCATTGTGTAGATGTA
CCACCTGGAGGAGTAGAGAAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCCATATAGG
TGGAACAACATTATGAGCTAACCAAGTACCCCGAGCTCTTGACTCTAGCTGCATATATATC
AAAAGATGGCCTAATCGGCCATCACTGGAAGAGAGGCCATTGGACTTGCAAACTTTAT
ATGCCCCAGTACAGGGGAATACCAGGGCCAAAAAGGGGAGTGGGTGGGCAGGGGAGTG
GGGGTGGGTGGATATGGGGGACTTTTGTATAGCATTGGAAATGTAAATGAGTTAAATAC
CTAATAAAAAATGGAAAAAAAAAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCAC
TTCACTGCTATCTCAACTTGCAACAGAAAGGGGAGTCACAAACCCAGGACAAACACAGT
GATTGAAGCGTCTTTGAATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCAT
TGTGAAAACCTTACGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGC
ATATTTGGCACCACTACCAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACA
CACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGG
CATCATCTGCTCACTCACCAACCAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTCTC
CTATTTTATGTCTAGGTTCCCTTCCCTGTTAGGAGATGGGAGAGGTCACGAAAGGAA
AGAATTTGTAGCCCCGTGAGCCAGCCCCGGGCCACAGAGCCTGCCACCAGACAGGAAAAGCC
CAGGGCTTACCAGCACAGGAGGAGCAAACTCGCAGGCGAGCCTGGGTGGCGCTGGTGGT
CCCGGGTCGATGGCCCGCCATTCCCAGAAGCCGAGGCTATAGCTGCGTCACCTGCCCGG
CCCTCCTCCCGAGTGAAAGACCCCTAGAGGCTGAGCAGACCCCAAGGCGGTGCAATTCCA
TTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCCTCGGGAAGTGAAGGGACCCAGAG

Exon1

AGTCTGGTAGATGTGGGAGCTGGGGTTTCAGGGCGAGACAGAGGGTGGGATGGGCAGAAAG
GGTCCAGGAAAAGGAAAGTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAA
CATCGCTTCAGTGACTGAAGCCAGGCCAAAAGGAGCGGGAAGGATTATATGTAGCCTGGGA
CGCTTTCATAAACACTGATGACGTGTTTGTGCAAGCAAGCAATTTGAGGAGAAACGGCT
GGGACGTCGGGAAAGAAAGGAGTGATCGATTAGTACTTGTAAAGTTTAGGTGAGTTGAGAAC
TAACCTAATCTATCTATTGAGGGAGAAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCA
GCAATCAGATAAAGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAAAAATAACA
GAAGCGCCTCCAGTTCTCTGAAGAGTCACTCCCCAGCTAGTGAAGACTAAGCCTACTAA
GCCTTTTGTCCCGTTGGAAGCAAGAACGTTCCCTCAATCAGGTGAAGGCTCTCCTCAGA

AGATTTCCTGCTCTGCTTATGTTACAAGAGGATTCAAAGCAAGACAGAAAGAGCTCAGG
TATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACATCATGTGGTA
TTCATTTACGGCTCTCTTCTTTATAACTAACTCTTAAGGTGCATATAGTCTCTTCTGTTTCC
CAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATC
AOCGAGAGAGTATTCAAAAATATTTCAGTGATGTAACAGTGACAGTGTAGGCATAGAAAGTA
ATCATTAGTAAATCTTAATTTGGGTAAACTCATTCAACAGCTCCAGGTGGGAGGGAT
CACTGAGCCTTCGCCACGTGCGGGTAAAGATATTTTCTAACAAGAGAAGCAGAATTCTTC
CTTGGCCATGCTCCCCATCACTGTGTCAAGTAAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAG
AGCAGACAGTGTATGCCTGCAAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTTACT
GTCTCCCGGTCAATTAGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGC
AGCATGCTAGAAAAGAAAATTGACCTTTTCTAGTGGGTATTACAGCTGTAAAAGTAT
TTTGGAAGGTAAAGCCAAATAAATAAAACACATATTAAATAATACAATGTTACAAAATT
GATCATATAAAGAAAGTACATTCAATAATGCAATGTGAAAAATATATATAATTTTATCTAT
TACTGGTGCAAGTTTTCTAAATTGCACATGTACTATTTTATATTTATAAAAAATATTTT
AAAATGTATATAAAAGTGTAAGAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAACT

↓ (BamHI)

TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCTGCTTTTCATTGGGA
AATGAGAAAGCTATGAAGATATGTTTACAATAATAAGCCATTAGTGATAAAGTCCAATG
GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCTGAGTCTATGCAAGTTTACA
CTCTAGGGAATAAGAGTTCTCTTCTCCAGATTTCACTAGCATTTGTTGTCACTATTATC
TTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAGGAATGTCAATTTGG
ATGCCCTGAACAATCTTTCAGGTCTTCTTTCAGTTCACTAGTCTATTCAATTATTGGATAA
TTGGGGGATGGTGTTAATTTTTTTCAGTTCTTATGGAATTCCAAAAACAAAAACAAAC
AAACAAACAAAAACCTCTGAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
AGAAATCTGCACAGAAATTTATTGCTACATTGTTTCAATTATTCAGACAGCCAAGAAATGTGGA
ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAAATGTGGAATGTGTACAAC
AGAGTCCCATGTGGCCATAAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT
GGAAATCAATTGGGCTAATCAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTT
CTCTGACAAACAGAAAGCTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATT
TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAACAAAAACA
AGCTAGCTGGGTGGGTAGGAGAGAAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGA
GGCACACGTTCAAACCAAGGGTGAAATCCAGCACAGAGAAGGGGAAGTAGACACAAAGT
TTCGCCACTAACCAAGAAGCCATTGTCAGTTGCTGCCTGCTGGGAGGGGCGTTCCAGTTT
CTCCAGTCTGACACTGTGTATAACAACAGTTGACAATAACAAAGTTGGCATGATGGATGG
TTTTTGTGCTATTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTTGGTGGTGTGTTGGTGGT
GCTGTGTTTTTCAATTTGTTCTTTTGAGAGAGAGAAGGAACATGAAATTGGGTGGGTAGGA
AGCTGGAACGATCTGGAAGAAGTTGGGGAAGAGAAAAATTTGTATGGAGCATATTTAA
ACAAACAAACAAACAAACAAAGGTTCAATTTGCCACAAAAAGGTGTGAATTAATTAAC
TAGTTACGACTCTTAAAGAAAAATATCCCAATTATCCAGAGTTGCTATGTATGCTGTGC
TAGGACTTTGCTTGAACCTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG - Artificial

NotI site

(SEQ ID NO:17)

Figure 48

CACAGACCTTCCTCTTCTAACTCTCTCCCCATCTTGTGCTTCATCCCAGACTTCAACAC
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCCTGTCAGGACAACAAGGAGGCTGAAG
GCAGACCCACACGTTTCCAAGTCTCCTGAGAGTCAATCCCCCTAGACTCATCTATAGCAG
GAAACCTGCTGTGATCTCCATTTCTTCTCTGACCACATCCCCAAGTTATCACAAGGAGTTT
TTCTCAAACCTTTCTCTCCAGCAAACCCCTTCAGCTCCTTGGGTACTTTCTCTAGCCCT
TCATTGGGAACCTGTGCTCCATCCAATGGATGGCTGTGAGCATCCACTTCTGTATAGAAT
CTTGGTCAGTGCAATCTTTGTATCCTCAAGAACACTGGGTCTGAAAAATTTAAACCCAAAG
AACTGTTTTTTGTTATGATTGCTGCAATCTCTTTCAATTCCAATAAAGAGTAAGCATCTCAT
TCCTTGTCTCCTCCTTTCAGTACCACCTGCTTTGTGCTCTTCTCAAAGAATCAATAAA
ACCAAAGTGATATAGATTTCATGGCAATTCCTCTAACTGCTACATCCACTCCAGTAGTATCTC
ACTTGGCAGGTGTAAAAGCCTGGAAGCAGTCAAGAGGCAGTTTCACAGAAACTTAGCCCTC
CTGGAACCTTGGCATTCCCATAGCTAGAATGCCAGATTGTCCCTGAGATATTGTGGTG
GGTCTTGCACTGCTTTCTTGCAGTATTTTACTGGATAAGAGTTAGAAATCTCAGGGCGAGCT
TAGCAAAAGTATACCTAGAAATCTTCATGACAGTCAGGTATTGCAAACTACATCTGCATTA
GAAGAAAGTTGGTAAATCTTCTGACAAATGGAGATTCCCTACAGATAACTTAAAAAGAAC
AGCTAAGTCACACTCATATGCAAGAATTTACCAAGGCCTAGGAAAGGGGGGGGGTACT
GCTTTATTTCATGATAAGGTCTGCTAGAGCAGAACCCCTGGGTGCTAGCTTTCAAGGTTTC
AAAGGTGTAGCATAAATTGTGACTAGAGTGTGAAATCTTTACCTGTCATTAGCTGACTCTA
GGCAGAGCTGTTTTATCTTTACTGTAAACATTACCTGGTTCTGTCAGTCTTTGAAGGCAT
TCCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCTGCTGTGACACCTACTCCTTTGT
TTTCTGTATTATATAAGCCTGGTGTTCCTTTGTGAAAAATTACATCCAGATACAGCACTC
CCTGTGTCTGTGTCTTTTGTCAATTTCTGGCCAACTCCATGCCCACTGCCAGAACCCCTA
GTCTTTCCACAGATTGAGGGAGGCCGACTGAGCCTGGTCCATGGCATCTAACCCTGTCA
GCTCACTGTTGGTGACTACCTCAAGGTACAAGCTCCATTACTAATGAAACAAAATTAGAT
AAGTGTGGGTCCAGGAAGCAGGTGTACACCCCTGTCTGAATGAACATTATGAAATGACTG
AAATAAGTTAAACCATCTCTTCTCGTTTGCTAATATAGCAAATAAACCGAGTTTCTGAGC
TGCTGCTGGTGTGTCTCCATCAGAGGGCAGAGCCAGTCTGATCCTAGCTTTCTGTATGTG
TGTCATTTGTTTCTTCAGTTCTGTTGGCCCATAGGAAATCCTAAGCCATGAAAGCCATG
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTGCCTTTCAACGTTTTGCCTG
TAAACAAGATATATGGTGGCAGTTTATAATCATAATAAGCTTTGAAATAATATATAACTC
CATTCTCATTCTGCTTCCACGCTGAGCATCTGTGTTTCCACAGGGACCACAAGAGCATTGGA
AAAGTAGTGATTTATGACCTGCTTTGTCTGTACTATAAAAGCTTCATGAAAGGGCAGCC
ATGTTGAAACATGGAACCTTGGGGTGACCTGTATCTGTGTTCTGGGTCTGCTCACTCATA
TTGTCTCCAGAATAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTTGTATAGCA
CGCCCGTGGAGTCCCAACATTCTACTTCTGTAAATCTGTATATGGTAGAAAAAGTTAATTT
ATGTGATTCTTCCAACCTCAAATATTTCAAATCTTTTAGCCCTCAGCCTGGGATTTCTTTG
ACTAAGTCTATTGATTTGGGAAGATCTCAGTGGTTAGGATTGCAGTTCATGATGTTTCATACG
TCAGGCTAAAGCTGAAAAATATGACAAATGAAATGTCAAATGTCAATGTGCTGGGAATGTG
AGTGTTAGGGGGTTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT
CTATCTAGTTTTGAATAATGTTCCCTAGTATACAGCAATTAATTTGAAATGAATAGCTTC
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCATGTTTTCATATTCTAAGGAT
AGGTCTCATTCAACCACATGTTTATCATTTGGGATCTGCATTTTCTGATGCAAAATGATTT
ATTCTTCCAGAGCACTGGAATTGGGTTGAATCATCTTATAACGGCCAAAACATAAATGCTTC
TGTGCTAAACAGAGAGTTACAAGACCTTTTTATGTGGATGGCAGCATTTTAGTCATCCTTA
TGACAGAAATGTGAGAGTGGAGCTCCCACTGGGGGAGGGGCTGGTCTTGGCAGGATCTC
TTGGAACATCACACAAAAGAAATTCAAATTATGAAATGCACATGATCCATCCAGAATGTG
ACTTTTGACTCTTGAACATGAGCTTTTAAAGTACGTTTGGCTGTTTCAAGCTTGACTTTGAG
GTGAAGGAAAGCTCGCCAACTCCTTTTTATATGTAACACAATATATCAAGATCTAATGTGA
GACAGTATGCCAGTCCCAAGATCTGTCAATGACTGAAGACACATTGGGATGTTATCACT
AAGGCAGGAGAAGGCAAGCTACAGTGAAGCCAGTTCACTATAAAGCTTTATGAGAAATT
AGATAAGAAGGGTTTCTAATTTTTAAATTTTTTTTATTAGATATTTCTTCATTTACATTT
AAATGCTATCCCAAAAGTCCCTATACTCCCCCTCACCTGCTCCCTACCCACCCACTC
CCATTTCTTGGCCCTGGTTACTTGTGATAGTGGTCAATGATCCACCAAGCTTTACATGCTC
ACTATCTGGTCTATTGCAAGAATGGGCTGCGAGCTGATGCAGTCAGATACAGACACCTAC
AGCCAAACAGTGAAGGAACCTTGGGGACTCTTATGGAAGAAAAGGAGGAAGGGTTATGG

009130 0042

2049530-03403

GCCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
AOCCTTGGGGGCTCTCAGAGTCTGAACCACCAACCATAGAACATTTCATGGGGCTGTACCCAG
GCCTCTCCACTCATATGTAACAGATATGTGGCTTGGCCTTCATCTGGGTCTGAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCACCTAGCTTTAGCCTT
GTAAAGACTTGAAGTTCTGAGGTGTGCGTGGAGGGTATACTCAGGGAGGCCCTCAOCTGC
TAGGAAGAGAAGAGGAGGGGGGAAGACTTGGGGAGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAAATAAAAAATAATTAATAAAAAAAAAAAGAATGGCTGCTGAGCCC
TACTCTAAAAACATTGCATCCCCCCCCCAATCATTTCAGTGACTACGAATTAAAAATCATT
GATACTAACAATAGATGTAGGAAACTATTGTAACTTCTTTGTGAECACGAGTGGTATTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTCTTTTCATTTCCCTGTACATGCATG
TAGCTTAATGATGTTTCAGTGAATTAATAAATAACAATGAAGAATAAAGACAACCTGTATTTT
AAGGATTCCTTCGTATATATTTAAAAATCTAAGGTGGTCAOCTGGAAGAAATGTCTTCAGT
TTTCTATATATGTTTACTCTATCGTATGTTAATTAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACCTTATAAGAAAGTAAAAACATTCATGAAATTGTGAAGGTTACTTTTC
CTTACCCCTCAGAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTTGAA
AGAATGTAATAACCTACTAAGGCAACATAGAAAGTTGAAATTAATAAGAGTAGACAGGA
GAAGTAATAAGGCAATAATGAATATTTGCTTTAAATAGTTCTTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAATAATAAACATTTCTACTTGAAAATTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGAACCAAGCAGCTTAT

(SEQ ID NO:18)

Figure 49

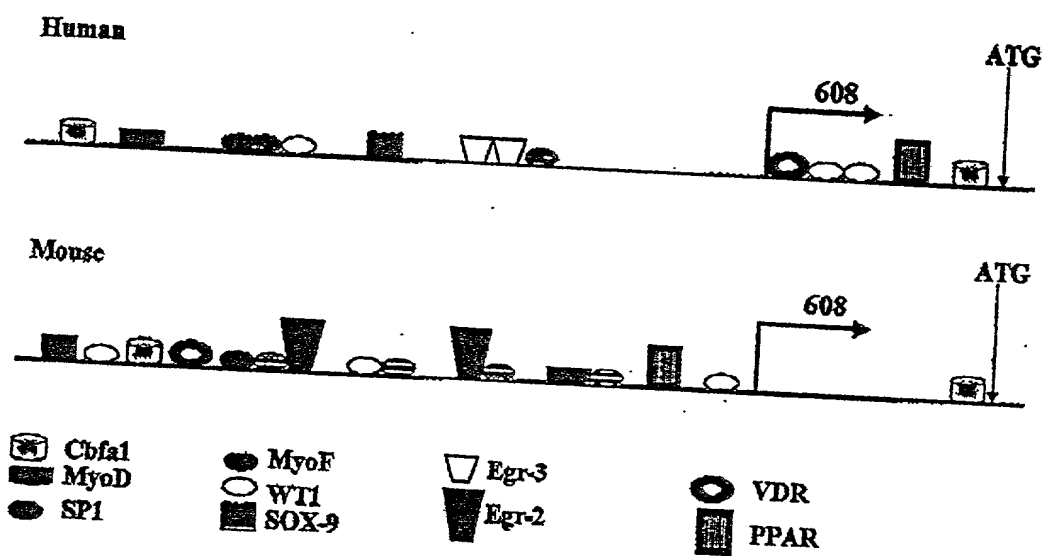


Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTTCCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACCTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTTCATTTGCGGCTCTCTTCTTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGAACCAACCCAGATGCCCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTA AAAACAACGACTTTCATAAAATTTTAA
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAACTGAGTCTTAAGGATTCATATTA ACTCCAGTTCCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTA ACTTCACCAGCTTCCTACACAGAGCA
GTTACTATCCTTGCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCAATTTAACCCCCCAA
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAATAAATAGG
AGAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTA CTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCTGTCACACAGGTCTCCAGGG
T TAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTCCCTTGATCATTAAATGACTAGA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GGCCACCTCAGCTTTGGCAGCTCCATTCTGTGTGGAGCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTT
CACAAGTCTTTTTAGAAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

095150 031550

TGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAA
TAAACTGACTTTAAAAAAGCTTATGATACTTGTACAGAGTAAATCTTCCATA
AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAACCTTTAGAAA
TTTTGTAAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
CAACAAACAACAACGACAACAAACCCATAAAAATTATCACGGCAAATGAACTAAGC
CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTCCCTTAATAGTCAA
ATATGCAGAAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
GTAGATTAGAGGACACAAAGCAACAACCTACAGAGTGAAGTAATCCAGAGACTTAAT
GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGT
ATCTTCATGTAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAA
GAGGAAGAGGATAGGTGCCTCTCCCAGGGCGGAAGAGTACATTCGAAAGTCAGACA
CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
CTTGACTCTAGCTGCATATATATCAAAGATGGCCTAATCGGCCATCACTGGAAAGA
GAGGCCATTGGACTTGCAAACCTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAA
AAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG
CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG
AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAAACCTA
CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
ACCACTACCAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
TCATCTGCTCACTCACCCACCCAAGCTTAATCCTTTTCCTTCCCTGCAATAGTACCTC
TCCTATTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC
AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAAACCTCGCAGGCGAGCCTGG
GTTGGCGCTGGTGGTCCCGGGTCGATGGCCCGCCCATTCCCAGAAGCCGAGGCTATA
GCTGCGTCACCTGCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
CCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTTACGGG
CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA
AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG
TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTGCGAAAGAAGGA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAATCAGATAA
AGGAAAGCTTTGGTTAGTTTGGAATGTATGATACCATTAAAATAACAGAAGCGCCT
CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
TGCTCCCGTTGGAAGCAAAGAACGTTCCCTCAATCAGGTGAAGGCTCTCCTCAGAAG
ATTTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAG
GTATTGCCAACTCTTTTGTTAAATACAGTTTGAGGCTTAAGTGTACGGGAACTCATG
TGGTATTCATTTACGGCTCTCTTCTCTTATACTAACTCTTAAGGTGCATATAGTCTC
TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
GCTTTTAAATCATCACGCAGAGAGTATTCAAAAATATTCAAGTGATGTAACAGTGACA
GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTAAACTCATTATAA
CAGCTCCAGGTTGGGAGGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
TCTAACAAAGAGAAGCAGAATTCCTTGGCCATGCTCCCCATCACTGTGTCAGTAA
GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCCTCCCGGTCATTAGTTGGCTCTG
AAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
GACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTATTTTGAAGGTTAAGCCA
AATAAATAAAACACATATTAAATAATACAATGTTACAAAATTTGATCATATAAAGA
AGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAAATATTTTTAAATG
TATATAAAAGTGTAAGAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAACTTTA
ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCA
ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
GTTTACACTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTTGTTGT
CATCATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTC
TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
TTCCAAAAAACAAAAACAAACCAACCAACCAAAAAACCTCTGAAACTAGAACTACC
AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
AACGGATAAAGAAAATGTGGAAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
GGGTGGGTAGGAGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

2019-11-03 10:00:00

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
ATGATGGATGGTTTTTGTGCTATTTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGGT
GGTTGTTGTGGTGGTGGCTGTGGTTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
ATGAAATTGGGTGGGTAGGAAGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAG
AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAGGTTCATTTT
GCCACAAAAAGGTGTGAATTAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
AATTATTCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAAGTGGCCCTA
TAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGAAGTCAGCTGCT
TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCAACCCCTGGGAGCAGGGTCTGTC
CTCGCCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTCTGGGACCTGA
CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
GACCTTGCCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCTGGTACTTTGCCCC
CCAAACACCTGTGATTACTTTTTAATAGTTTTCTTCTAAAATGGGTTCATACAAACCTT
ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTTGAAGCC
CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
AAATGAAAAAGATACAGTTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
TGTCATTGAATATTTTAAAAAAGCATTCTTTTGCAATTTATAAATACCCATTACAAA
ATGGCTTACTTAAATACTTGCCTTACTAAATCTGACAAATTATGGTGATATTTTGAA
GGTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTACGTTTGGAG
GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
GGGCTGAGCACTTTGCAATGAGCATTGTGCTCATTGTGCTGGCAGGTTTTATGATAAC
TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTCAACTGAGAACATGCCTGA
AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTCTTAATTAGTGATTCATGGG
GCAGGGCCCAGTCCATTGTTCTGTTGTTACCATTTCTCAGGCCTATTAAAAAAAAAAAA
AACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
TGCATCAGCTCCTGCCTCCAGGTTCTACCTATTTGAGTTCCTGTCCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTT
GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
GTTTGTTCATTTTGGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
ATTTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTACAGAGTCAGTGACA
AGACCTTCTCGGGCTTGCAGTCCTTGCAGGTGAGATAGGTAGAGGGTGATGGAGGC
TGAGAAGAGAGGTGCAACTGTGGGTTATACCCAAAAGCTGCTGATTCCCGTGGGAG
ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCATTGT
AATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATGCTGGGGAAAATGCTTAC

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTTCAAGATCCCAGAGACT
CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACCTTGAC
TCAGAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
GCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAAGGAAGGAAGGAAGGAAG
GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
AGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
GAAAGGGGGAGAAAGAAGAGAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACCTGAAACTACAATTATT
CCCACAAATGAAGAAAATGCTGTAAGAGTTTTTCACACACCGAAGTGGGAAACTTAA
GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG
AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
CTCAACTCTTTAAGCAAACAATGCTGTTGTTTCATGAAAAGCACATAAAGTACATAT
GTCCCATATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
CACCCACTGTTACAGAGACTTTAAACATGAAACTGGAACCTATGTCTAGTGTTTTGA
CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTTCATGTATCCTGCATGAAATTC
TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGTTTG
AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTTATTAGTATTTTCATGTTCTTT
TCTTTTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGA
AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAATTT
TTTGCCTTAATTCAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAG
ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACA
GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTCGGCTACATG
ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
TTAAAAAATAAACTATTGATGTGGAATGATCTTAATTTGGGGAAAAGACAATT
TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAATG
AGATGGTAGAAAGGAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCCAAAGTTCATGGTCAATGTGTGT
GGAGACACCATAATAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAATCTAA
AAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
GGAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTAACTGCTTTTTA
CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTCT
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGCAATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACCTATCTTGTTTGTGTTTCTTTTTTT
ATARKACGTATTTTCCTCAATTTCAATTTAGAATGATATCCCAAAGTCCCCCATAACC
TCCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGCGTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

CG5430 : protein

Figure 51

MPKRAHWGALS VV LILL WGHPR VALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
RINLGFNSIQALSETSFAGLTKLELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
LSTIRHLYLAENMVRTLPASMLRNMPLENLYLQGNPWTCDCEMRWFLWDAKSRGIL
KCKKDKAYEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSEEEQEQ
EEDGGSQ LILEKFQLPQWSISLNM TDEHGNMVNLVCDIKKPM DVYKIHNLQTDPPDIDIN
ATVALDFECPMTRENYEKLWKLIA YYSEVPVKLHRELMLSKDPRVSYQYRQDADEEAL
YYTG VRAQILAEP EWVMQPSIDIQLNRRQSTAKKVLLSYTQYSQTISTKDTRQARGRS
WVMIEPSGA VQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPD SKFSIL
SSGWLRIKSMEPSD SGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDTVTIGKNPGES
VTLP CNALAIPEAHL SWILPNRRIINDLANTSHVYMLPNGTLSIPKVQVSDSGYYRCVAV
NQQGADHFTV GITVTKKGSGLPSKRGRRP GAKALSRVREDIVEDEGGSGMGDEENTSR
RLLHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR
INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSL SLEVTPPFPAVSPPSASP
VQTVTS AEESADVPLL GEEHVLTGISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
SEKTEEITSTEGDLKGTAAPT LISEPYEPSPTLHTLDTVYEKP THEETATEGWSAADVGSS
PEPTSSEYEPPLDAVSLAESEPMQYFDPDLET KSQPD EDKMKEDTFAHLTPPTI WVND S
STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKKG MKEMSQTLQGGNM
LEGDP THSRSS ESEGQESK SITLPDSTLGIMSSMSPVKKPAETT VGTLLDKD TTTVTTPR
QKVAPSSTMSTHPSRRRPNGRRRLRPNKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
QVESSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGT PPRKHGKRPNKHYTPSTVSSRA
SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDMYMTTTRKIYSSYPKVQETLP
VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
GTPTWNPSRTAQPGRLQTDIPVTTSGENLTDPLLKELEDVDFTSEFLSSLTVSTPFHQEE
AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSPS
TILMSLGQTTTTK PALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEG TQHMSPNE
LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM
SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSNGRLPFFT NKTL SF
PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLD FGPPAPPLLHTPQT TGSPSTN
LQNIPMV SSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
TAETD TVFPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
QYMCTASNLHGLDRMVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
SWIFPDRR VWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
AALPPVIHQEKLENISLPPGLSIHIHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTSPRRTD VRYG
GTLKLDCSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLV VKSVTDK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

THE

[illegible]

TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTTACGG
TGGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
AGACGCCAG GTGCAAAGGC TCTTTCCAGA GTCAGAGAAG ACATCGTGGA
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
GAAGCATTCTG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA GGTCGCAGAG
TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
GGGCACAGAA GTACCCCCGT TGATTAAAC CACAAGTCCT CCATCCTTGA
GCCTAGAAGT CACACCACCT TTTCCTGCTG TTTCTCCCCC CTCAGCATCT
CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCTCA GCCAGCATGG
GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
AGATGTTGGA TCGTCACCAG AGCCACATC CAGTGAGTAT GAGCCTCCAT
TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTAAAT GACTCCAGTA
CATCACAGTT ATTTGAGGAT TCTACTATAG GGAACCAGG TGTCCCAGGC
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
TCGAAGGAGA CCCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
ACCGGCACAA GCAAACCCCA CCCACAACCTT TTGCCCCATC AGAGACTTTT
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
GAGTTCTCTG GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
AACATAGAAA CATTGTTACT CCCAGTTCAG AAACCTATACT TTTGCCTAGA
ACTGTTTCTC TGAAAACCTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC
AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
 TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
 GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
 AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
 GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
 GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
 TTTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCACC
 AGGAAGAAGC TGGTTCCTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
 GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC
 CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA
 CTGCTGCCCC GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
 ATGTCTTTGG GACAAACCAC CACCCTAAG CCAGCACTTC CCAGTCCAAG
 AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTT TTGAATTATG
 TGGGGAATCC AGAAACAGAA GCAACCCAG TCAACAATGA AGGAACACAG
 CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
 ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
 GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
 GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
 AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
 TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
 TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
 ATCAAGTACA ACAATTCCTC TCCCATTGCA CATGTCCAAA CCCAGCATTC
 CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTAATCCAAA
 GTGTTTGGA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC
 TCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
 CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTCACCCG GAGACCCAG
 ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
 TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC
 CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCTCA
 ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
 CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
 CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
 GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
 TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
 TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
 AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
 GTTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
 ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT
 CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
 CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTCCT
 GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GGCGTCTATA AGTGCGTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCTG CACGTGGCGG CACTGCCCCC
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
GCATTACAT TCACTGCACT GCCAAGGCTG CGCCCCCTGCC CAGCGTGCGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
CCAAGGACAG CGGGCGCTAT GAGTGCGTGG CCGCCAACT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TGCGTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA
GACGGGAGTC TGGTGAATC CTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTTAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCACCTCTC TTATTAGAA AGCCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATT CAGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
CCCGAGGGTG TGGTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCAAT
GGCACCAGTC TGCAGAGTGG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTTCAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACTT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTTGTA
AGGGAAGCCA GGTGGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTCATCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCAGAT TGACAAGTCA
TCTTTCAGTT ATTTCTCTG TCACTTCAAA ACTCCAGCTT GCCCAATAAG
GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
TACATACATA CAGCTACCAT TTTATATGAA AAAAGAAAAA CATTTCTTCC
TGGAACTCAC TTTTATATA ATGTTTTATA TATATATTTT TTCCTTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTTAAATTA
ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTTT AAAAAATTTT TCTCCAACCT CCTTCAAATT CAGTCACCAC
TGTTATATTA CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
AGATTTCTTT GTATGCAAAG TTTTGTGGA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAACTTC TCTGCAGTAT CTGGCTTGTC
CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
CCATGAATAA TACACGACCT GTTATTTCCA TGAAGCTTTT ACTGTATTTT
TAAGGTCAAT AACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAA AAAAA

Figure 53

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC
CCTGGGGGCAAGGCCTGTCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTTCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC
ATTCACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGACAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACCTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
ACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGGGCTCAACTTTCTCCGCCTGGTGCCTTTGGAA
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC
TCTTTCATTAAGTTCTTATACCTTGTCTGATAACTTCCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGATACAGGAGAAGCCAGATGTAATAAATGCAAAAAAGATAGAAGTCCCTCTAGTGCTCAG
CAGTGTCCACTTTGATGAACCTTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCT
TTCCAGTGTGCCAAGCCAACCTTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGT
AGTCTGCTTTTATCTCTCCCCAAGGTTTCATGGCACCCCTTTGGCTCCCTCACTTTGAATATGACAGAT
CAGTCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCCATTGCATTC
ACTGAAGAAAATGACTACATCGTCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAGATTAC
GGTCACATTGAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAAGGAGC
CACTTGCTTAGTGAAACACCCGACGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATT
TTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTTGCAG
CTGAACAGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAACTACTTTA
CCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTG
GAACATACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCAGGCCAAGGAGACCCACCCACAC
GTGGATTGGCTTCTAGCTGATGGAAAGTAAAGTGAGAGCCCTTATGTGAGTGGAGATGGACGGATCCTA
ATAGACAAAAGTGGAAAATTGGAACTCCAGATGGCTGATAGTTTGGACACAGGCGTATATCACTGTATA
AGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTGCAGGCC
TATCAGGAAAATGGGATTCTACACAGTTTTTCATTGGTGAAACACTTGATCTTCCATGCCATTCTACT
GGTATCCAGATGCCTCTATTAGCTGGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGAC
AAGAAAGTTCTAAACAATGGCACATTAAGAATATTACAGGTACCCCGAAAGACCAAGGTATTATTCGC
TGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTTGTATTTTCCAAGTTTCAGTCAAGATGAAAGGACAA
AGGCCCTTGGAGCATGATGGAGAAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTT
AAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAAACACACCTCA
AGCACAAGTAAGAGGCACAACATTCGGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGT
TTTAGGGAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCAGT
TTGGAGAAAAGCTAAAAAGAAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCACCCCAAGT
GTCACCCAACTCCCAAACATACCTGGTGAAGAAGACGATTCCCTCAGGCATGCTCGCTCTACATGAGGAA
TTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAACAATA
TCTGATAGTCTCTATGACAAACATAAATTATGGCACAGAATTCTCTCTGTTGTGAATTCACAAATACTA
CCACCTGAAGAACCCACAGATTTCAAACCTGTCTACTGCTATTAAACTACAGCCATGTCAAAGAATATA
AACCACACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTTCATCCACTGTCTTTCCACTGCTA
CTTGGAGCAACTGAATTTTCAGGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCC
CCAATAACAGTAAGGACTATGATCAAAGATGTCAATGTCAAAATGCTTAGTAGCACCACCAACAACTA
TTATTAGAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCAC
AATCACTTCTATTCTCACACTACTCAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCT
GCTCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACAATCCCGCTGTTTCAGACGCTTTGGGAGG
CAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGA
TACAGCATTTTCAGGTCAACAACAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTG
CTCAATGTGACATGTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTT
CCAAGTGTGCTCCCATCACCTTCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACCT
CTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAAGAGAAAACACACCCACAAATAAAA
TATTTTCAGGACTGAAATTTCCCAAGTGACTCCAACCTGGTGCAGTCATGACATATGCTCCAACATCCATA
CCCATGGAAAAAATTCACAAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGA
GATTCAGTGATTACATCGTCACTTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACA
AGGTTTTCAGAAAGGAAAATTCCTTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTA
AGGAATCAACATAAAGTTAGTTTACAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCCTGCTTTTA
CCACAGAGACAAAGTTCCCTTTTCATTTTACCACACTTTTCAACAAGTGTGATGCAAAATTCATCTAAT
ACCTTGACTACCGCTCACCACACTACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAG
CTTCCCTTCCACCCCTTAACCTATGCTTCTTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATC
ATATCAACGCAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCTACTTATGAAACC
CAAACAGAGAGATCTAGAGCACAACAATACAAAGAGAACAGGAGCCTCAAAGAAGAACAGGACTGAC
CCAAACATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCTGCTCTGGCA

TTCACTCATTCCCCACCAGAAAACACAACCTGGGATTTCAGCACAATCAGTTTTTCATTCAAGAACTCTT
 AATCTGACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCT
 GAAACAACCTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCATTAGACACTCCCATA
 CCACCATTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCCATCTCCCCTCCCCTTTACTCAGAGAGCA
 GTTACTGACACACGTGGCGACTCCCATTTCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCC
 TCAAGGCACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTA
 CATTCTACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACCTCCATCTCCCTGG
 GCAGAATACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCCAAAAGCCAGAAGTAAGC
 ATGTTGGCTACTACAGGCTGTCCGAGGCCACCACTCTTGTTCAGATTGGGATGGACAGAAGAACACA
 AAGAAGAGTGACTTTTGATAAGAAACAGTTCAAGAAGCAACAACCTCCAACTCCCTTCCCTTTGACTCT
 TTGTCTAGGTATATATTTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTACTATTCCAGCT
 AACTCAGATGCCTTTCTTCCCTGTGAAGCTGTGTGAAATCCCCTGCCACCATTCACTGGACCAGAGTT
 TCAGGACTTGATTTATCTAGAGGAAAACAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCTGTCC
 ATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCGCATCCAATCTGTTGGCAGACAG
 CACCTTCATGTCACCTGTCTGTGGTTTCTATCCTCCCAGGATCCTGGAGAGACGTACCAAGAGATC
 ACAGTTCATTCCGGAAGCACTGTGGAAGTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
 TGGATTCTTGCAAAACCAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGGTGACGGTT
 GACGGAACATTGGTCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAAC
 CCAGGTGGCCAGGATTCACTGCTGGTTAAAATACAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAA
 AGGAGGCAAGTCATTGTAGGCACCTGGGGTGAAGTTTAAACTGCCCTGTACTGCAAAAGGAACTCCT
 CAGCCCAGCGTTTACTGGGTCTCTGATGGCAGTGAAGTGAACCAATTACAGTTTACCAATTCCAAG
 TTGTTCTTATTTTCAAAATGGGACTTTGTATATAAGAAACCTAGCTCTTCAGACAGGGGCACCTTATGAA
 TGCATTGCTACCAGTTCCACTGGTTCCGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACC
 AGCCCCAGGATAGAAGCTGCATCCAGAAAAGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAAC
 TGCTCAGCCACTGGGGAGCCCCAAACCCCAATAATGTGGAGGTTACCATCCAAGGCTGTGGTCCGACCAG
 TGGAGCTGGATCCACGTCTACCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGT
 GTCTACTTGTGTGTGGCAAGAAACAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTG
 AAACCTGCCAAAATTGACCACAAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTA
 GATTGCAAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAAC
 AATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATTGGAATTTA
 TACTTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGTCTATGCCAGAACACCCCTAGGGAAA
 GATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATAAGGCAGAGTAACAAAACCAAC
 AAGAGAATCAAAGCTGGAGACACAGCTGTCTTACTGTGAGGTCACTGGGGATCCCCAAACCAAAATA
 TTTTGGTTGCTGCCTTCCAATGACATGATTTCTTCTCCATTGATAGGTACACATTTTCATGCCAATGGG
 TCTTTGACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCCGAAATCCCAGT
 GGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACA
 AACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCTGAAGGG
 ACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTTCTCACAGCCCCATACTATGGAAGC
 AGAATCACAGTCCATAAAAATGGAACCTTGGAAATTAGGAATGTGAGGCTTTCAGATTACGCCGACTTT
 ATCTGTGTGGCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTG
 AGAAGACCGACATTTAGAAATCCATTAAATGAAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTG
 AATTGCTCTGTTGATGGTAACCCACCACTGAAATAATCTGGATTTTACCAAATGGCACACGATTTTCC
 AATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAACAACTCGG
 GAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTA
 GAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAGATCT
 CTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTAT
 GTAGTAGACAGGCCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAA
 GCAACAGCTTATGACAGAGGAACTATATCTGTAAGGCTCAAATAGTGTGGTCATACACTGATTACT
 GTTCCAGTAATGATTGTAGCTTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTACCAGG
 ACAGGGGCAGCCTTTTCACTGCTGCTGCTGGCCTTGGGAGTTCCCAAGCCAGAAATCACATGGGAGATG
 CCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGAAGTGAGCAGCTTCACTTACAA
 GGTACCCTAGTCATTGAGAATCCCCAAACCTCCGATTCTGGGATATACAAATGCACAGCAAAGAACCCA
 CTTGGTAGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 54

ORIGIN

1	MKVKGGRG	ITC	LLVSFAVICL	VATPGGKACP	RRCACYMPTE	VHCTFRYLTS
51	IPDSIPP	NVE	RINLGYN	SLV	RLMETDFSGL	TKLELLMLHS
101	FSDLQALQVL	KMSYNKVRKL	QKDTFYGLRS	LTRLHMDHNN	IEFINPEV	FY
151	GLNFLRLVHL	EGNQLTKLHP	DTFVSLSYLQ	IFKISFIKFL	YLSDNFLTSL	
201	PQEMVSYMPD	LDSL	YLHGNP	WTCDC	HLKWL	SDWIOQEKPDV
251	SAQQCPLCMN	PRTSGKPLA	MVSAAAFQCA	KPTIDSS	SLKS	KSLTILEDSS
301	SAFIS	PQGF	APF	GS	LT	LNMTDQSGNEANM
351	YIVLNTSFST	FLVCNIDYGH	IQPVWQILAL	YSDS	PLILER	SHLLSETPQL
401	YYKYKQVAPK	PEDIFTNIEA	DLRADPSWLM	QDQISLQ	LN	R
451	QYSSDAQITL	PRAEMRPVKH	KWTMISRDNN	TKLEHTVLV	G	GTVGLNCPGQ
501	GDPTPHVDWL	LADGSKVRAP	YVSEDGRILI	DKSGKLELQ	M	ADSFDTGVYH
551	CISSNYDDAD	ILTYRITVVE	PLVEAYQENG	IHHTVFIGET		LDLPCHSTGI
601	PDASISWVIP	GNNVLYQSSR	DKKVLNNGTL	RILQVTPKDQ		GYRRCVAANP
651	SGVDFLIFQV	SVKMKGQRP	L	EHDGETEGSG	LDES	NPIAHL
701	SALMEAEVGK	HTSSTS	KRHN	YRELT	LQRRG	DSTHRRFREN
751	IDPQHWAALL	EKAKKNAMPD	KRENTTVSPP	PVVTQLPNIP		GEEDDSSGML
801	ALHEEFMVPA	TKALNLPART	VTADSR	TISD	SPMTNIN	YGT
851	LPPEEPTDFK	LSTA	IKTTAM	SKNINPT	MSS	QIQGTTNQHS
901	TEFQDSDQMG	RGREHFQSRP	PITVRTMIKD	VNVKML	SSTT	NKLLLESVNT
951	TNSHQTSVRE	VSEPRHNH	FY	SHTTQILSTS	TFPSDPHTAA	HSQFPIPRNS
1001	TVNIPLFRRF	GRQRKIGGRG	RIISPYRTPV	LRRHRS	SIFR	STTRGSSEKS
1051	TTAFSATVLN	VTCLSC	LPRE	RLTTATAALS	FPSAAPITFP	KADIARVPSE
1101	ESTTLVQNPL	LLLENKPSVE	KTTPTIKYFR	TEISQVTPTG		AVMTYAPTSI
1151	PMEKTHKUNA	SYPRVSS	TNE	AKRDSVITSS	LSGAITKPPM	TIIAITRFSR
1201	RKIPWQQNFV	NNHNPKGR	LR	NQHKVSLQKS	TAVMLPKTSP	ALPQRQSSPF
1251	HFTTLSTSV	QIPSN	TLTA	HHTTTKTHNP	GSLPTKKELP	FPPLNPMPLS
1301	IISKDSSTKS	IISTQTAIPA	TTPTFPASVI	TYETQTERS	SR	AQTIQREQEP
1351	QKKNR	TD	PNI	SPDQSSGFTT	PTAMTPPALA	FTHSPPENTT
1401	RTLN	LDVIE	ELAQASTQTL	KSTIAS	ETTL	SSKSHQSTTT
1451	PFLSSSATLM	PVPISPPFTQ	RAVTDTRGDS	HFR	MTNTVV	KLHESSRHNL
1501	QMPSSQLEPL	TSSTSNLLHS	TPMPALTTVK	SQNSKLTPSP		WAEBYQFWHKP
1551	YSDIAEKGGK	PEVSM	LATG	LSEATTLVSD	WDGQKNTKKS	DFDKKPVQEA
1601	TTSKLLPFDS	LSRYIFEKPR	IVGGKAASFT	IPANS	DAFLP	CEAVGNPLPT
1651	IHWTRVSGLD	LSRGNQNSRV	QVLPNGT	LSI	QRVEIQDRGQ	YLCSASNLF
1701	TDHLHVTLSV	VSYP	PRILER	RTKEITVHSG	STVELK	CRAE
1751	LANQTVVSES	SQGS	RQAVVT	VDGYL	VJANL	SYNDRG
1801	LLVKIQVIAA	PPVILEQRRQ	VIVGTWGESL	KLPCTAKGTP		QPSVYVWLSD
1851	GTEVKPLQFT	NSKLFLFSNG	TLYIRNLASS	DRGT	YECIAT	SSTGSERRVV
1901	MLTMEERVTS	PRIEASQKR	TEVNFGDKLL	LNCSATGEPK		PQIMWRLPSK
1951	AVVDQWSWIH	VYPNGSLFIG	SVTEKDSGVY	LCVARNKMGD		DLILMHVSLR
2001	LKPAKIDHKQ	YFRKQVLHGK	DFQVDC	KASG	SPVPEISWSL	PDGTMINNAM
2051	QADDSGHRTR	RYTLFNNGTL	YFNKVGVAEE	GDYTCYAQNT		LGKDEMKVHL
2101	TVITAAPRIR	QSNKTNKRIK	AGDTAVLDCE	VTGDPKPKIF		WLLPSNDMIS
2151	FSIDRYTFHA	NGSLTINKVK	LLDSGEYVCV	ARNPSGDDTK		MYKLDVVSKF
2201	PLINGLYTNR	TVIKATAVRH	SKKHFD	CRAE	GTPSPEVMWI	MPDNI
2251	YYGSRITVHK	NGTLEIRNVR	LSDSAD	FICV	ARNEGGS	VL
2301	RRPTFRNP	PN	EKIVAQLGKS	TALNCSVDGN	PPPEIIWILP	NGTRFSNGPQ
2351	SYQYL	IASNG	SFIISKTTRE	DAGKYRCAAR	NKVG	YIEKLV
2401	LTYAPGTVKG	ISGESLSLHC	VSDGIPKPNI	KWTMPSGYV	V	DRPQINGKYI
2451	LHDNGTLVIK	EATAYDRGNY	ICKAQNSVGH	TLITVPMIV		AYPPRITNRP
2501	PRSI	VRTGA	AFQLHCV	ALG	VPKPEITWEM	PDHSL
2551	HLQGT	LVION	PQTS	DSGIYK	CTAKNPLGSD	YAATYIQVI*

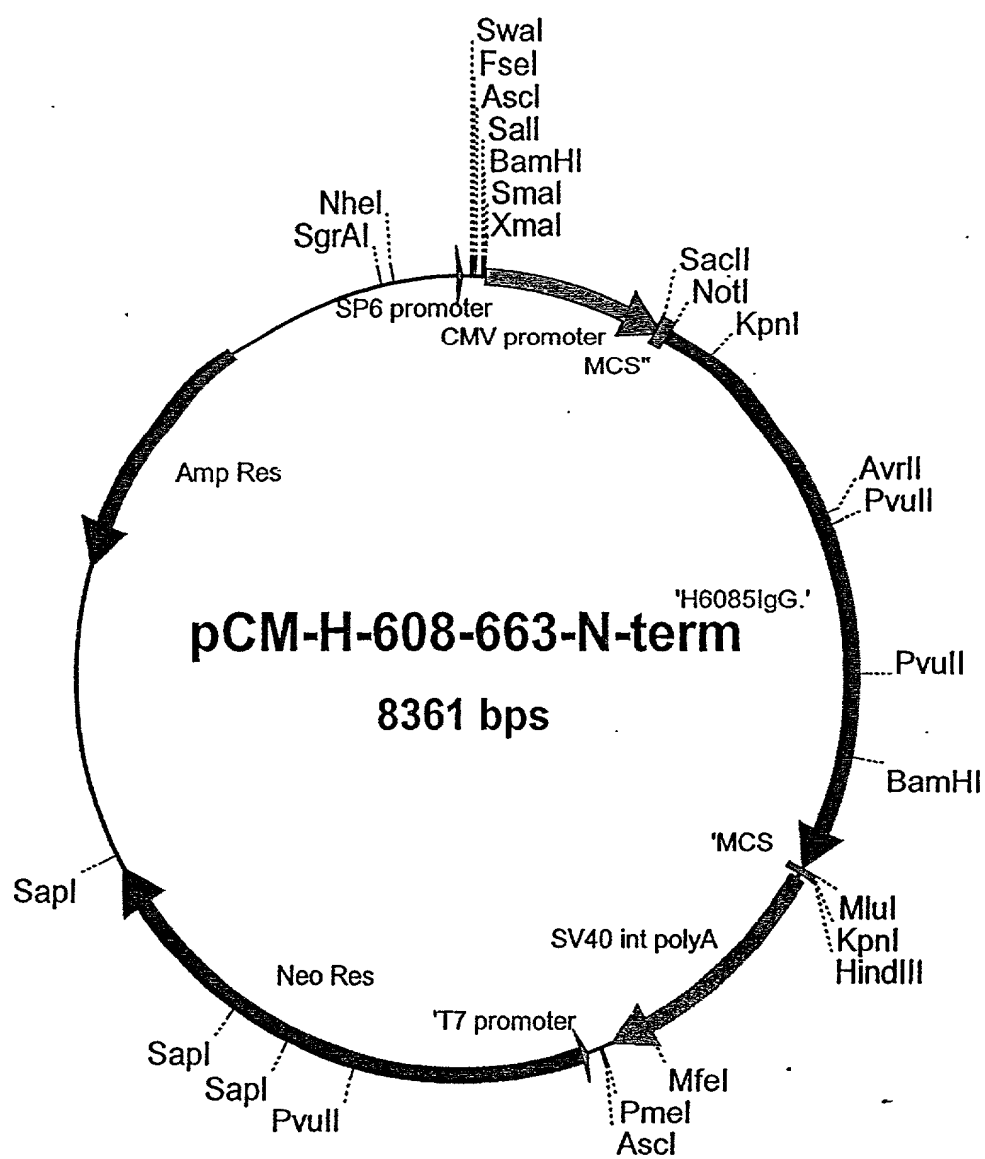
bioRxiv preprint doi: <https://doi.org/10.1101/2019.04.10.344402>; this version posted April 10, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

MQVRGREVSGLLISLTAVCLVVTGPSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNSLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVLSYLO
IFKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPCLMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNG
SASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFST
NLVCSVDYNIHQPVWQLLALYSDSLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEA
DVRADPPWFQOEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWMTMILMMNN
PKLERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELOM
ADSFADAGLYHCISTNDADADVLTyrITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKQDQGHYQCVAAANPSGADFSSFKV
SVQ

Figure 56

pCM-H-608-663-N-term



rat -TSKHLPPYSLPKTLKKPRIIGGKAASFTVPANSIVFLPCEAVGDPLPIIHWTRVSSGX
human_5+3_corrected TTSKLLPFDSLRYLFEKPRIVGGKAASFTIPANSDAFLPCEAVGNPLPIIHWTRVS-GL
mouse_5_corrected -----

rat KISQGTQKSRPHVLNGLTSLIQRVSIQDRGQYLCSAFNPLGVDHFHVSLSVVFPARILD
human_5+3_corrected DLSRGNQNSRVQVLNGLTSLIQRVEIQDRGQYLCSASNLFGTDLHLVLSVVSYPPRIILE
mouse_5_corrected -----

rat RHVKRITVHFGSTVELKCRVEGMFRPTVSWILANQTVVSETAKGSRKVVVTPDGTLLIYN
human_5+3_corrected RRTKEITVHSGSTVELKCRAGEGRPSPTVTNILLANQTVVSESSQGSROAVVTVDGTLVLHN
mouse_5_corrected -----

rat LSLYDRGFYKCVASNPGGQDSLLVKIQVITAPPVIEQKROAIVGVLGGSXLPLCTAKGT
human_5+3_corrected LSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRRQVIVGTWGESLKLPLCTAKGT
mouse_5_corrected -----

rat PQPSVHWVLYDGTTELKPLQLTHSRFELYPMGTLYIRSIAPSVRGTYECIATSSSGSERRV
human_5+3_corrected PQPSVYVWVLSDGTVEVKELQFTNSKLFNGLTYIRNLASSDRGTYECIATSSSGSERRV
mouse_5_corrected -----

rat VILTVEEGETIPRIETASQKTEVNLGEKLLNCSATGDPKFRILWRPLPSKAVIDQWHEM
human_5+3_corrected VMLTMEERVTSPIEASQKTEVNFGLKLLNCSATGEPKQIMWPLPSKAVVDQ-----
mouse_5_corrected -----

rat GSRIEHVIENGSLVVGSVTEKDAGDYLCVARNKMGDDILMHVRLRLTPAKIEQKQYFKQ
human_5+3_corrected GSWIHVYPNGSLFVIGSVTEKDSGYLVCVARNKMGDDILMHVSLRLKPAKIDKQYFQKQ
mouse_5_corrected -----

rat VLHGKDFQVDCRAGSGSPVPEVSWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGLTYFNV
human_5+3_corrected VLHGKDFQVDCRAGSGSPVPEISWSLPGDTNINNAQADDSGHRTRYTLFNNGLTYFNKV
mouse_5_corrected -----

rat GMAEEGDYICSAQNTLGKDEMKVHLTVLTAIPRIQSYKTMRLRAGETAVIDDCEVTGEP
human_5+3_corrected GVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIQSNKTKRIKAGDTAVIDDCEVTGDP
mouse_5_corrected -----

rat KPNVFWLPSNNVISFSNDRFTFHANRLSIHKVKPLDSGDYVCVAQNPFGDDTKYKLD
human_5+3_corrected KPPIFWLPSNDMISFSIDRYTFHANGSLTINKVKLLDSGEYVCVARNPFGDDTKMYKLD
mouse_5_corrected -----

rat IVSKPFLINGLYANKTVIKATAIRHKKYFDCRADGIPSSQVWIMPNGNIFLPAPYFGSR
human_5+3_corrected VVSKPFLINGLYTNRTVIKATAVRHKKHFDCAEGTPSPPEVMWIMPNDIILTAPYFGSR
mouse_5_corrected -----

rat VTVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLMLRRPTFRNPFNKIVIA
human_5+3_corrected ITVHPNGTLEIRNVRLSDSADFTCVARNEGGESVLVVQLEVLMLRRPTFRNPFNKIVIA
mouse_5_corrected -----

rat QAGKPVALNCSVDGNPPPEITWILPDGTQFANRPMSPYLMAGNGSLILYKATRNKSGKY
human_5+3_corrected QLGKSTALNCSVDGNPPPEITWILPNGTRFSNGPQSYQYLIASNGSFTIISKTTREDAGKY
mouse_5_corrected -----

rat RCAARNKVGYTEKLILLEIGQKPVILTYEPGMVKSVSSEPLSLHCVS DGIKPNVKNWTP
human_5+3_corrected RCAARNKVGYTEKLIVLEIGQKPVILTYAPGTVKSIGESLSLHCVS DGIKPNIKWIMP
mouse_5_corrected -----

